AF/1638

FULBRIGHT & JAWORSKI L.L.P.
A REGISTERED LIMITED LIABILITY PARTNERSHIP
600 CONGRESS AVENUE, SUITE 2400
AUSTIN, TEXAS 78701-3271
WWW.FULBRIGHT.COM

RHANSON@FULBRIGHT.COM
DIRECT DIAL: (512) 536-3085

TELEPHONE: (512) 474-5201
FACSIMILE: (512) 536-4598

June 28, 2005

**CERTIFICATE OF MAILING**
**37 C.F.R. 1.8**

I hereby certify that this correspondence is being deposited with the U.S. Postal Service with sufficient postage as First Class Mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on the date below:

June 28, 2005 — Date
Robert E. Hanson

**Mail Stop Appeal Brief-Patents**
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Re: *SN 08/113,561 "METHODS AND COMPOSITIONS FOR THE PRODUCTION OF STABLY TRANSFORMED, FERTILE MONOCOT PLANTS AND CELLS THEREOF" by Thomas R. Adams, et al.;*
*Our Ref. DEKM:055US; Client Ref. 51719 US 02*

Commissioner:

Transmitted herewith for filing are:

1. An Appeal Brief (an original and two copies);
2. A check for $500.00 to cover the Appeal Brief filing fee; and
3. A return postcard to acknowledge receipt of these materials. Please date stamp and mail this postcard.

If the check is inadvertently omitted, or the amount is insufficient, or should any additional fees under 37 C.F.R. §§ 1.16 to 1.21 be required for any reason relating to the enclosed materials, or should an overpayment be included herein, the Commissioner is authorized to deduct or credit said fees from or to Fulbright & Jaworski L.L.P. Account No.: 50-1212/DEKM:055US.

Respectfully submitted,

Robert E. Hanson
Reg. No. 42,628

REH/vv
Enclosures

PATENT

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re Application of:
Thomas R. Adams *et al.*

Serial No.: 08/113,561

Filed: August 25, 1993

For: METHODS AND COMPOSITIONS FOR THE PRODUCTION OF STABLY TRANSFORMED, FERTILE MONOCOT PLANTS AND CELLS THEREOF

Group Art Unit: 1638

Examiner: Fox, David T.

Atty. Dkt. No.: DEKM:055US

CERTIFICATE OF MAILING
37 C.F.R. §1.8

I hereby certify that this correspondence is being deposited with the U.S. Postal Service as First Class Mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on the date below:

June 28, 2005
Date

Robert E. Hanson

# BRIEF ON APPEAL

## TABLE OF CONTENTS

PATENT

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | |
|---|---|
| In re Application of:<br>Thomas R. Adams *et al.* | Group Art Unit: 1638 |
| Serial No.: 08/113,561 | Examiner: Fox, David T. |
| Filed: August 25, 1993 | Atty. Dkt. No.: DEKM:055US |
| For: METHODS AND COMPOSITIONS FOR THE PRODUCTION OF STABLY TRANSFORMED, FERTILE MONOCOT PLANTS AND CELLS THEREOF | |

## BRIEF ON APPEAL

**Mail Stop Appeal Brief - Patents**
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Appellants hereby submit an original and two copies of this Appeal Brief. The date for filing this Brief is June 28, 2005. The fees for filing this Appeal Brief are attached. However, should any additional fees become due under 37 C.F.R. §§ 1.16 to 1.21 for any reason relating to the enclosed materials, or should an overpayment be made, the Commissioner is authorized to deduct or credit said fees from or to Fulbright & Jaworski Deposit Account No. 50-1212/DEKM:055US.

Please date stamp and return the attached postcard as evidence of receipt.

07/05/2005 BABRAHA1 00000076 08113561
01 FC:1402 500.00 OP

## I. REAL PARTY IN INTEREST

The real party in interest is Monsanto Company, the parent company of assignee DeKalb Genetics Corp.

## II. RELATED APPEALS AND INTERFERENCES

There are no interferences or appeals for related cases.

## III. STATUS OF THE CLAIMS

Claims 1-68 were filed. Claims 1, 5-66 and 68 were canceled. Claims 2-4 and 67 are therefore currently pending and are the subject of this appeal. A copy of the appealed claims is attached as Appendix 1.

## IV. STATUS OF AMENDMENTS

No amendments were made subsequent to the Final Office Action.

## V. SUMMARY OF CLAIMED SUBJECT MATTER

The invention relates to genetically transformed monocotyledonous plants. Specification at page 3, lines 10-13. More particularly, it relates to fertile, transgenic maize plants transformed with a DNA sequence encoding a fatty acid desaturase gene, wherein the DNA sequence is capable of being transmitted to subsequent plant progeny and is expressed. Specification at page 306. Expression of the fatty acid desaturase yields plants with altered seed oil properties. Specification at page 45, lines 18-19.

## VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

(A) Are claims 2-4 and 67 properly rejected under 35 U.S.C. §112, first paragraph, as failing to comply with the written description requirement?

(B) Are claims 2-4 and 67 properly rejected under 35 U.S.C. §112, first paragraph, as not being enabled by the specification?

Appellants note that the Final Office Action rejected claims 2 and 3 as indefinite under 35 U.S.C. §112, second paragraph, for depending upon a canceled claim. Appellants intend to correct the error by amendment upon the allowance of the case or reopening of prosecution and thus are not appealing the rejection.

## VII. ARGUMENT

### A. The Claims Meet The Written Requirement Under 35 U.S.C. §112, First Paragraph

The Examiner asserts that the claims lack an adequate written description under 35 U.S.C. §112, first paragraph, on the basis that the fatty acid desaturase genes incorporated into the claimed plants are not adequately described. For example, it was stated that claims drawn to maize plants transformed with a particular gene are inadequately described if the starting material, namely the gene, is itself inadequately described. Action dated May 13, 2004 at p.4. *Eli Lilly* was cited in this regard for the proposition that a claimed invention must be defined by a precise definition, such as by structure, formula, *etc.*, and MPEP §2163, p.156 was cited for the principle that a biomolecule cannot be defined merely by function when the function is not correlated with a structure. *Id.* at p. 4-5. In the Final Office Action the Examiner added a citation to the *University of Rochester* district court case for the holding that method claims are properly subjected to a written description rejection if the starting material required by the method is inadequately described. *University of Rochester v. G.D. Searle & Co., Inc.*, 249 F.

Supp. 2d 216; 68 U.S.P.Q.2D 1424 (D.N.Y, 2003). Final Action at p. 3-4. Finally, the Examiner asserted that desaturases were not known, stating that many of the earlier-submitted references were published after August 1993. Final Action at p. 4. As explained below, none of these arguments properly supports the rejection and thus the rejection should be reversed.

**1. The Rejection is Legally Unsupported**

The cases cited by the Examiner involve situations in which written was found lacking because the point of novelty was not described. In contrast, the fatty acid desaturase genes alleged here to have not been described were known in the art. The Examiner has nonetheless examined the claims as if they were directed to fatty acid desaturase genes *per se*. This position is contrary to the cited cases and well settled precedent holding that the specification need not disclose what is well-known to those skilled in the art and ***preferably omits*** what is well-known and already available to the public. *See Hybritech, Inc. v. Monoclonal Antibodies, Inc.*, 802 F.2d 1367, 1384, 231 U.S.P.Q. 81, 94 (Fed. Cir. 1986).

The distinction between the current situation and the cited cases is illustrated by comparing the facts at issue. In *Eli Lilly*, for example, the subject patent claimed a novel human insulin-encoding cDNA sequence but disclosed only a rodent sequence. A lack of written description was found because the specification failed to describe the human sequence being claimed. *The Regents of The University of California v. Eli Lilly and Co.*, 119 F.3d 1559, 1568; 43 USPQ2d 1398, 1405 (Fed. Cir. 1997). In *Amgen v. Chugai*, cited by the Examiner on page 5 of the Office Action dated May 13, 2004, the issue was what constituted conception of an invention directed to isolated DNA sequences encoding human erythropoietin. *Amgen Inc. v. Chugai Pharmaceutical Co.*, 927 F.2d 1200, 18 U.S.P.Q.2d 1016 (Fed. Cir. 1991). The Federal Circuit rejected an accused infringer's accusation of prior invention under 35 U.S.C. §102(g) based on conception of a generalized approach for screening a DNA library because the

methodology was not a definite and permanent idea of the complete and operative invention without knowledge of the structure of the gene sequence. *Id.* at 1206. The issue was therefore conception of a novel nucleic acid sequence, not description of a known nucleic acid sequence.

The Final Office Action attempted to supplement the foregoing legal shortcomings of the written description rejection by citing two additional cases, the *University of Rochester* district court case of March 2003, Order No. 00-CV06161L dated March 5, 2003, and the *Bayer v. Housey* Federal Circuit case. 340 F.3d 1367, 68 U.S.P.Q.2d 1001 (Fed. Cir. 2003); Final Office Action at p. 3-5. Neither of these cases is on point to the current situation as well.

At issue in the *University of Rochester* case were claims directed to a method of selectively inhibiting the enzyme COX-2 by administering a non-steroidal compound that selectively inhibits activity of the COX-2 gene product. The district court found that the patent at issue was invalid for failure to comply with the written description requirement because the applicants did not disclose a non-steroidal compound that selectively inhibits COX-2 and provided no specific a suggestion how it could be made. 249 F. Supp. 2d at 224. The case was taken on appeal and the Federal Circuit affirmed. *University of Rochester v. G.D. Searle & Co.,* 358 F.3d 916, 69 U.S.P.Q.2d 1886 (Fed. Cir., Feb. 13, 2004). The Federal Circuit noted in particular that the patent was invalid for written description because the required compound, *e.g.*, an inhibitor of COX-2, was not disclosed in the application and there was ***no pre-existing awareness in the art*** of a compound exhibiting the claimed activity. *Id. at 927.* The court emphasized that what was not described or known was what in fact was essential to the claimed invention - a compound that inhibits COX-2 – and that the inventors had neither possession nor knowledge of such a compound. *Id.* This authority therefore does not relate to the situation where a known compound is used to make a new claimed product. *Rochester* and the

corresponding line of authority are therefore inapposite to the current situation and provide no support for the rejection.

The Final Office Action also cited *Bayer v. Housey*, 340 F.3d 1367, 68 U.S.P.Q.2d 1001 (Fed. Cir. 2003) for the proposition that "processes of identification and generation of data are not steps in the manufacture of a final [drug] product." Final Action at p. 4. However, the issue decided in this case and referenced with regard to steps in the manufacture of a drug product was patent infringement under 35 U.S.C. §271(g), not written description. The case turned on the meaning of "product" under 35 U.S.C. 271(g) and whether this covered the importation of information gained from patented drug screening assays. *Id.* at 1371. The case therefore has no relevance to the current written description rejection and provides no support for the rejections made.

In sum, no legal basis has been provided for maintaining the written description rejection. The rejection made is directly contrary to well settled legal precedent holding that what is known in the art need not be described with particularity and is in fact preferably omitted from the specification. *See, e.g., See Hybritech, Inc. v. Monoclonal Antibodies, Inc.*, 802 F.2d 1367, 1384, 231 U.S.P.Q. 81, 94 (Fed. Cir. 1986). Findings of fact and conclusions of law by the U.S. Patent and Trademark Office must be made in accordance with the Administrative Procedure Act ("APA"). 5 U.S.C. § 706(A), (E), 1994; *see also In re Zurko*, 258 F.3d 1379, 59 USPQ2d 1693 (Fed. Cir. 2001). An Examiner's position on Appeal must be supported by "substantial evidence" within the record pursuant to the APA in order to be upheld by the Board of Patent Appeals and Interferences. *See In re Gartside*, 203 F.3d 1305, 1315, 53 USPQ2d 1769, 1775 (Fed. Cir. 2000). As the current rejections are unsupported in fact or law, the standards of the APA have not been met. Reversal of the rejection is thus respectfully requested.

## 2. Fatty Acid Desaturases Were Well Known in the Art

Unsupported conclusions to the contrary in the Final Office Action notwithstanding, numerous fatty acid desaturases were known and found in the literature prior to the August, 1993 filing date. Among these, McDonough *et al.* (**Exhibit A**) ("Specificity of unsaturated fatty acid-regulated expression of the *Saccharomyces cerevisiae* OLE1 gene."; *J Biol Chem.* 1992 Mar 25;267(9):5931-6) describe a *Saccharomyces cerevisiae* OLE1 gene encoding delta-9 fatty acid desaturase, an enzyme which forms the monounsaturated palmitoleic (16:1) and oleic (18:1) fatty acids from palmitoyl (16:0) or stearoyl (18:0) CoA. Fox *et al.* (**Exhibit B**) ("Stearoyl-acyl carrier protein delta 9 desaturase from *Ricinus communis* is a diiron-oxo protein." *Proc Natl Acad Sci U S A.* 1993 Mar 15;90(6):2486-90) describe a gene encoding a stearoyl-acyl carrier protein delta 9 desaturase from castor that was expressed in *Escherichia coli.* The authors compared the primary structures of catalytically diverse proteins to identify conserved amino acid motifs involved in eukaryotic fatty acid desaturation.

Reddy *et al.* (Abstract - **Exhibit C**) ("Isolation of a delta 6-desaturase gene from the cyanobacterium *Synechocystis* sp. strain PCC 6803 by gain-of-function expression in *Anabaena* sp. strain PCC 7120" *Plant Mol Biol.* 1993 May;22(2):293-300) describe the cloning of a delta 6-desaturase from the cyanobacteria *Synechocystis* that is responsible for the conversion of linoleic acid (18:2) to gamma-linolenic acid (18:3 gamma). A delta 12-desaturase gene linked to the delta 6-desaturase gene was also identified and expression of the delta 6- and delta 12-desaturases in *Synechococcus* deficient in both desaturases carried out to result in the production of linoleic acid and gamma-linolenic acid. Arondel *et al.* (Abstract - **Exhibit D**) ("Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis.*" *Science.* 1992 Nov 20;258(5086):1353-5) describe a gene from *Arabidopsis thaliana* that encodes an omega-3 desaturase. Transgenic tissues of both mutant and wild-type plants of the model dicotyledonous

plant *Arabidopsis thaliana* were found to have significantly increased amounts of the fatty acid produced by this desaturase. PCT Application Publ. No. WO 91/13972 describes plant Δ9 desaturases (**Exhibit E**), European Patent Application Publ. No. EP 0616644 describes soybean and *Brassica* Δ15 desaturases (**Exhibit F**), and European Patent Application Publ. No. 0537178 describes soybean stearoyl-ACP desaturases (**Exhibit G**).

Appellants therefore have shown that numerous examples of fatty acid desaturases were found in the literature and available to the public before the August 25, 1993 filing date. For example, each of the references submitted as Exhibits A-G were published before August 25, 1993. These examples demonstrate that genes encoding fatty acid desaturases were well known in the art.

The specification itself further describes in detail how such fatty acid desaturase genes alter grain composition traits. For example, it is taught that genes may be introduced to alter the balance of fatty acids present in seed oil providing a more healthful or nutritive feedstuff, and may be used to block expression of enzymes involved in fatty acid biosynthesis to alter proportions of fatty acids present. As explained, changes in oil properties may be achieved by altering the type, level, or lipid arrangement of the fatty acids present in the oil. Among representative catalytic steps mentioned for modification include the desaturations from stearic to oleic acid and oleic to linolenic acid resulting in the respective accumulations of stearic and oleic acids.

These examples demonstrate that genes encoding fatty acid desaturases were well known in the art and that the specification fully describes their use in altering grain composition traits. What was not known in the prior art was that they could be expressed for benefit in maize. The inventors have overcome this deficiency and for the first time describe methods enabling the

expression of desaturases to alter maize grain composition traits. No assertion has been made that the transformation of maize and transgenic maize plants generally have not been described. The specification contains numerous descriptions of transgenic plants and working examples showing the introduction of transgenes into plants.

The specification, for example, indicates after Table 8 that fertile plants were obtained from 267 different transgenic lines produced. In Table 9, the specification describes the creation of numerous transgenic maize plants with a variety of different genes using many different regulatory elements. For example, the table shows the creation of R0 transgenic plants and confirmation of transgene expression in these plants and progeny using the following genes: a *uidA* reporter gene, a *bar* selectable marker gene conferring herbicide tolerance, a *hyg* gene conferring resistance to hygromycin, an *aroA* gene conferring tolerance to the herbicide glyphosate, a *Bacillus thuringiensis* endotoxin gene, and a Z10 altered seed storage protein. The Table further shows that transgenic maize callus was obtained transformed with a C1 anthocyanin pigmentation gene, a *lux* luciferase reporter gene, potato and tomato *pinII* proteinase inhibitor genes conferring insect resistance, an *mtlD* protein conferring enhanced stress resistance and a *deh* gene conferring resistance to dalapon herbicide. While an actual reduction to practice for fatty acid desaturase genes is not described, it is well settled that Appellants need not have done so. This is underscored by the numerous working examples in the specification and detailed teachings in the specification fully establishing possession of the invention.

In conclusion, Appellants have affirmatively established on the record a written description for the claimed subject matter and demonstrated the lack of any legal basis for doubting the sufficiency of the description. Reversal of the rejection is thus respectfully requested.

## B. The Claims Are Enabled

The Examiner rejected claims 2-4 and 67-68 under 35 U.S.C. §112, first paragraph, as not enabled for fatty acid desaturase genes or expression of the genes in plants. For example, the Examiner asserted that evidence that desaturase genes were well known in the art was non-persuasive by citing *Genentech Inc. v. Novo Nordisk A/S*, 108 F.3d 1361, 42 USPQ2d 1001 (Fed. Cir. 1997) for the proposition that the specification, not the knowledge of one skilled in the art, must supply the enabling aspects of the invention. Declaratory evidence submitted by Appellants showing that desaturase genes successfully express in maize plants was disregarded as "insufficient to demonstrate that the specification enabled the claimed invention." As explained below, the rejection is unsupported and should be reversed.

### 1. The Rejection is Legally Unsupported

Appellants first note that the Action does not contest the fact that the specification fully enables transformation of maize with heterologous genes. Appellants further note that the claims are directed to maize plants transformed with desaturase genes, not fatty acid desaturase genes *per se*, as these sequences are known. The authority cited in the Final Action does not relate to such a situation. For example, in *Genentech* the subject patent claimed a method of producing hGH hormone using a cleavable fusion expression. *Id.* at 1365. A lack of enablement was found because the specification did "not describe in any detail how to make hGH using cleavable fusion expression." *Id.* The court agreed that the specification need not disclose what is already well known in the art, holding that it is "the specification, not the knowledge of one skilled in the art, that must supply the ***novel*** aspects of an invention in order to constitute adequate enablement." *Id.* at 1366 (emphasis added). In the current application, what is ***novel*** is a

composition of maize plants transformed with desaturase genes, not the known desaturases themselves. The *Genentech* case is therefore inapposite to this situation.

### 2. Appellants Have Affirmatively Established the Enablement of the Claims

Appellants have further presented affirmative evidence demonstrating enablement in the form of the Declaration Dr. Virginia Ursin. **Exhibit H**. Dr. Ursin describes studies showing that the expression of Δ6 and Δ15 desaturases in maize results in an alteration in the fatty acid profile of corresponding transgenic plants that renders them identifiable over the corresponding non-transgenic plants. *Id.* at ¶6-7. As explained, the results showed that the ***two*** desaturases were expressed and that alteration of fatty acid profiles in maize occurs in a predictable manner that is consistent with the enzymatic activity of the fatty acid desaturase that is introduced. *Id.* at ¶7. This evidence therefore establishes that expression of a fatty acid desaturase in maize would in fact occur in a predictable manner that distinguishes transgenic plants from corresponding non-transgenic plants.

The Final Action dismissed this evidence as non-persuasive because Dr. Ursin used a transformation technique "which is not the technique disclosed by Applicant," and because Dr. Ursin "utilized fatty acid desaturase genes which were disclosed... well after the effective filing date." Final Action at p. 6. These statements, which Appellants take as true for the purposes of this argument only, were used as the basis of the Examiner's conclusion that the Declaration "is insufficient to demonstrate that the specification enabled the claimed invention." However, the statements made by the Examiner do not justify maintenance of the rejection. First, the Examiner has not contested the enablement of the application for transformation of maize plants and the transformation method is irrelevant to whether a fatty acid desaturase gene is expressed. The goal is to introduce a foreign gene and, the specification having already enabled this, the

method used is irrelevant. Whether *Agrobacterium*-mediated transformation or microprojectile bombardment was used is thus irrelevant.

With regard to the public availability of the desaturase genes used by Dr. Ursin, this also does not negative enablement. While the Δ6 and Δ15 genes expressed in maize were not the same as the fatty acid desaturase genes disclosed in Exhibits A-G, the studies show that fatty acid desaturases are expressed consistent with their known enzymatic characteristics in transgenic maize. Further, other Δ6 and Δ15 desaturases were known as shown in Exhibits C and F. As explained by Dr. Ursin, the studies disclosed in the Declaration:

> demonstrated that expression of a fatty acid desaturase gene in maize alters the fatty acid profile in a manner that renders the transgenic plants identifiable over corresponding non-transgenic plants. The results further confirm that the alteration of fatty acid profiles in maize occurs in a predictable manner that is consistent with the enzymatic activity of the fatty acid desaturase that is introduced into a given maize plant.

The Declaration therefore establishes that fatty acid desaturase gene expression in maize occurs in a predictable manner and serves to distinguish transgenic plants from corresponding non-transgenic plants. Furthermore, with regard to the asserted non-availability of desaturases, it has already been shown above that each of the fatty acid desaturases disclosed in Exhibits A-G were available as of August 25, 1993, and thus can be relied upon consistent with *In re Glass*.

Finally, the Examiner attempted to support the foregoing rejection by stating that Appellants in the bottom paragraph of page 9 of the Response to Office Action dated October 13, 2004 acknowledged the unpredictability of the claimed invention. A review of the Response, however, demonstrates that it refers only to the teachings of the cited prior art, not the predictability of the invention in view of the specification teaching as it is relevant to enablement. As explained in the Response, the cited prior art was entirely prophetic with regard

to creation of a single transgenic maize plant. In contrast and as illustrated further in §VII.B.3 below, the specification reports production of fertile transgenic plants from 267 different transgenic lines, and reports the production of progeny plants containing and expressing numerous foreign genes. One of skill in the art would have therefore been without any reasonable expectation in arriving at the inventions based on the prior art, but would have been fully enabled for practice of the claimed invention upon possession of the specification.

### 3. The Working Examples Demonstrate Enablement of the Claims

The specification contains working examples demonstrating the production of transgenic plants from numerous different transgenes and demonstrates confirmation of the expression of these transgenes. For example, after Table 8 the specification discloses that fertile transgenic plants were obtained from 267 different transgenic lines. In Table 9, the specification describes the creation of transgenic maize plants with a variety of different genes using different regulatory elements. For example, Table 9 shows the creation of R0 transgenic plants and progeny in which transgene presence ***and expression*** have been confirmed for a diverse collection of transgenes including: a *uidA* reporter gene, a *bar* selectable marker gene conferring herbicide tolerance, a *hyg* gene conferring resistance to hygromycin, an *aroA* gene conferring tolerance to the herbicide glyphosate, a *Bacillus thuringiensis* endotoxin gene, and a Z10 altered seed storage protein. The Table further shows that transgenic maize callus was obtained transformed with a C1 anthocyanin pigmentation gene, a *lux* luciferase reporter gene, potato and tomato *pinII* proteinase inhibitor genes conferring insect resistance, an *mtlD* protein conferring enhanced stress resistance and a *deh* gene conferring resistance to dalapon herbicide. These examples coupled with the evidence presented above fully demonstrate the enablement of the claims for transgenic expression of fatty acid desaturases.

In conclusion, Appellants have affirmatively presented evidence on the record establishing enablement and at the same time demonstrated the lack of any legal basis for rejecting the claims. Reversal of the rejection is thus respectfully requested.

## VIII. CONCLUSION

It is respectfully submitted, in light of the above, that none of the claims are properly rejected. Therefore, Appellants request that the Board reverse the pending grounds for rejection.

Respectfully submitted,

Robert E. Hanson
Reg. No. 42,628
Attorney for Appellants

FULBRIGHT & JAWORSKI L.L.P
600 Congress Avenue, Suite 2400
Austin, Texas 78701
(512) 536-3085

Date: June 28, 2005

## APPENDIX 1: LISTING OF APPEALED CLAIMS

1. (Canceled)

2. (Previously amended) Cells obtained from the plant of claim 67, wherein said cells comprise the DNA composition.

3. (Previously amended) Progeny of the plant of claim 67, wherein said progeny comprise the DNA composition.

4. (Previously amended) Seeds obtained from the plant of claim 3, wherein said seeds comprise the DNA composition.

5-66. (Canceled)

67. (Previously amended) A fertile, transgenic maize plant, the genome of which has been augmented by the introduction of a DNA composition comprising a gene encoding a grain composition trait comprising a fatty acid desaturase gene so that the transgenic plant exhibits one or more phenotypic characteristics that render it identifiable over the corresponding untransformed maize plant which does not comprise said gene, and wherein said gene is transmittable through normal sexual reproduction of the transgenic maize plant to subsequent generation plants.

68. (Canceled)

## APPENDIX 2: EVIDENCE APPENDIX

Exhibit A: McDonough *et al.* (*J Biol Chem.* 1992 Mar 25;267(9):5931-6); submitted with Response to Office Action Dated October 13, 2004 and entered by Final Office Action

Exhibit B: Fox *et al.* (*Proc Natl Acad Sci U S A.* 1993 Mar 15;90(6):2486-90); submitted with Response to Office Action Dated October 13, 2004 and entered by Final Office Action

Exhibit C: Reddy *et al.* (*Plant Mol Biol.* 1993 May;22(2):293-300); submitted with Response to Office Action Dated October 13, 2004 and entered by Final Office Action

Exhibit D: Arondel *et al.* (*Science.* 1992 Nov 20;258(5086):1353-5); submitted with Response to Office Action Dated October 13, 2004 and entered by Final Office Action

Exhibit E: PCT Application Publ. No. WO 91/13972; submitted with Response to Office Action Dated October 13, 2004 and entered by Final Office Action

Exhibit F: European Patent Application Publ. No. EP 0616644; submitted with Response to Office Action Dated October 13, 2004 and entered by Final Office Action

Exhibit G: European Patent Application Publ. No. 0537178; submitted with Response to Office Action Dated October 13, 2004 and entered by Final Office Action

Exhibit H: Declaration of Dr. Virginia Ursin; submitted with Response to Office Action Dated October 13, 2004 and entered by Final Office Action.

# EXHIBIT A

THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1992 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 267, No. 9, Issue of March 25, pp. 5931-5936, 1992
Printed in U.S.A.

# Specificity of Unsaturated Fatty Acid-regulated Expression of the *Saccharomyces cerevisiae OLE1* Gene*

(Received for publication, November 5, 1991)

**Virginia M. McDonough‡§, Joseph E. Stukey‡¶, and Charles E. Martin‖**

*From the Bureau of Biological Research, Nelson Biological Laboratory, Rutgers University, Piscataway, New Jersey 08855-1059*

**The *Saccharomyces cerevisiae OLE1* gene encodes the Δ-9 fatty acid desaturase, an enzyme which forms the monounsaturated palmitoleic (16:1) and oleic (18:1) fatty acids from palmitoyl (16:0) or stearoyl (18:0) CoA. Previous studies demonstrated that *OLE1* mRNA levels and desaturase enzyme activity are repressed when either 16:1 Δ-9 and 18:1 Δ-9 are added to the growth medium (1). The polyunsaturate, linoleic acid (18:2, Δ-9,12), which is not a product of the enzyme, is also a strong repressor. The specificity of the *OLE1* transcriptional regulatory sensor was examined by testing the response of *OLE1* promoter-*lacZ* fusion reporter genes to fatty acids that differ in chain length, degree of unsaturation and double bond positions. Monounsaturated and polyunsaturated fatty acids that contain a Δ-9 double bond are strong repressors of reporter gene activity and native *OLE1* mRNA levels. Monounsaturated fatty acids containing double bonds in the Δ-10, Δ-11, or Δ-5 positions showed no repression of reporter enzyme levels although they were rapidly incorporated into membrane lipids and some supported growth of an *OLE1* gene disrupted strain. Although 17:1 Δ-10 does not repress *OLE1* transcription, lipid analysis showed that it replaces almost all of the endogenous 16:1 Δ-9 and 18:1 Δ-9 in cellular lipids and *OLE1* mRNA levels are strongly repressed. This suggests that additional systems regulate desaturase activity by post-transcriptional mechanisms that differ from the transcriptional sensor in their responses to specific fatty acids.**

In *Saccharomyces cerevisiae* unsaturated fatty acids are formed by the Δ-9 fatty acid desaturase, which introduces a double bond between carbons 9 and 10 of palmitoyl (16:0)-[1] or stearoyl (18:0)-CoA to form palmitoleic (16:1) or oleic (18:1) acid. The desaturase, which is encoded by the *OLE1* gene, appears to be a major determinant of cellular membrane and storage lipid composition, and its importance for normal cell growth is suggested by the fact that these monounsaturates can comprise greater than 70% of the total cellular fatty acids. Since unsaturated fatty acids are essential for membrane expansion in growing cells and are major components of storage lipids in stationary phase cells, the enzyme must be regulated in response to an array of metabolic and physiological stimuli. In order to regulate the enzyme, cells must be able to discriminate between saturated and unsaturated acyl species in pools of precursors or in complex lipids.

One component of desaturase regulation in yeast involves a response to fatty acids that are added to the growth medium. When the monounsaturated 16:1 Δ-9 and 18:1 Δ-9 products of the enzyme are added to wild type yeast cultures they are rapidly incorporated into cells and assimilated into membrane lipids. Under those conditions, *OLE1* mRNA levels are sharply reduced and enzyme activity is reduced to undetectable levels (1).

Although *Saccharomyces* does not form polyunsaturates under normal growth conditions, linoleic acid (18:2 Δ-9,12) is also a strong repressor of desaturase mRNA levels and enzyme activity (1). It is preferentially incorporated into membrane lipids of wild type cells and when added to the growth medium will replace almost all of the naturally occurring monounsaturated fatty acid population after several generations of growth (2, 3).

The association of unsaturated fatty acid repression with reductions in *OLE1* mRNA levels, suggests that transcriptional controls are a major regulatory component of desaturase activity. It is not clear, however, how cells might detect the presence of the fed unsaturated fatty acids and regulate *OLE1* expression. Given the initial observation that both mono- and polyunsaturated fatty acids can trigger *OLE1* repression, one possibility is that this mode of regulation is a part of a system that monitors and maintains membrane lipid unsaturated fatty acids at levels required for cell growth and other functions. Tests of yeast *ole1* mutants have also shown that a variety of other unsaturated species (see Ref. 4 for review), including polyunsaturated acids, can fulfill the cellular requirement for unsaturated fatty acids. Given this observation, we examined the range of specificity of the *OLE1* regulatory sensor as a step toward determining the molecular basis of the response and the relationship between that mode of desaturase regulation and the regulation of the gene under other physiological conditions. Measurements of *OLE1* mRNA levels by themselves do not allow one to differentiate between transcriptional and post-transcriptional control mechanisms. We therefore measured the activity of *OLE1* promoter-*lacZ* fusions to assess the contributions of transcriptional and other modes of regulation.

* This work was supported by a grant from the Charles and Johanna Busch Memorial Fund from the Bureau of Biological Research and by National Science Foundation Grant 84-17802. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Charles and Johanna Busch Predoctoral Fellow.

§ Supported by the Anne B. and James H. Leathem Scholarship Fund.

¶ Present address: Dept. of Biology, Princeton University, Princeton, NJ 08544.

‖ To whom correspondence should be addressed: Bureau of Biological Research, Nelson Biological Laboratory, Rutgers University, Busch Campus, P.O. Box 1059, Piscataway, NJ 08855-1059. Tel.: 908-932-4081 or 908-873-2752; Fax: 908-932-5870.

[1] The abbreviations used are: *x*:*y* Δ-*n*, fatty acyl groups containing *x* carbon atoms and *y* cis double bonds located at position *n* relative to the carboxyl or carbonyl end of the hydrocarbon chain; bp, base pair(s).

## MATERIALS AND METHODS

*Strains, Growth Medium, and Transformations*—Haploid strains L8-25A (*MATα, OLE1, ura3-52, leu2-3, leu2-112, his-4*) and L8-14C (*MATa, ole1Δ::LEU2,leu2-3,leu2-112,ura3-52,his4*) used in this study were derived from the diploid JSY67X (2). Yeast cells containing the *lacZ* fusion plasmids p40 or pCT:OLE were grown at 30 °C on uracil dropout SD medium (5) supplemented with 1% tergitol (SDT) and 1 mM of the appropriate fatty acid (obtained from Sigma or Nu-Chek Prep). All fatty acids used in this report contained double bonds in the cis configuration. Recombinant DNA manipulations using *E. coli* are described by Maniatis (6) and Ausubel *et al.* (5).

*Nucleic Acid Blots*—RNA blots using total cellular yeast RNA were carried out as described previously (1). Either the L32 ribosomal subunit mRNA (7) or the yeast actin gene was used as an internal control. DNA probes were made either by the random primer method or by chemiluminescent modifications (BRL) under conditions recommended by the manufacturer.

*β-Galactosidase Assays*—β-Galactosidase assays of cells containing plasmids with intact *OLE1* promoter sequences were performed essentially as described (5). Cell densities for those assays were determined either by measurement at $A_{600}$ or by hemocytometer counts. Assays of extracts from cells with plasmids containing hybrid *OLE1* upstream promoter elements and *CYC1* TATA and downstream sequences were done using the procedure of Lue and Kornberg (8). Those assays were correlated with the total protein in cell extracts determined by the method of Bradford (9) using the Bio-Rad assay kit.

*Repression/Derepression Studies*—Cells containing an *OLE1-lacZ* reporter gene were used to test unsaturated fatty acid regulation in two ways. First, fatty acids were tested for their ability to repress reporter gene activity in cells in an initially derepressed state. Cells were grown overnight on medium without unsaturated fatty acids, then washed with ice-cold distilled water. Aliquots were resuspended in 25 ml of fresh minimal medium or in a medium containing a specific unsaturated fatty acid species. The final cell density was less than $1 \times 10^7$/ml. After 8–10 h of growth, cells were harvested, washed in ice-cold distilled water 3 times, and then tested for β-galactosidase activity. Activities were compared with those from cultures grown in medium containing no unsaturated fatty acids and those grown in repressing medium containing either 1 mM 18:1 or 18:2. Secondly, fatty acids were tested for their ability to maintain the fully repressed condition. Cells were grown in the presence of 1 mM 18:1 Δ-9 or 18:2 Δ-9,12 overnight. Low cell densities were maintained so that the medium was not depleted of the fed unsaturated fatty acid ($<8 \times 10^6$/ml). Cells were washed as described above, and aliquots were resuspended in minimal growth medium or media containing unsaturated fatty acids at 1 mM concentrations. After 4–6 h of growth, cells were tested for β-galactosidase activity and the results were compared with the previously described controls.

*Lipid Extraction and Analysis*—Total cellular fatty acids were obtained by HCl-methanolysis of extensively washed cell pellets according to the procedure of Browse and Somerville (10) and direct extraction of methyl esters in hexane ether (1:1) or by saponification of washed cell pellets (11), followed by petroleum ether extraction, acidification, and re-extraction of fatty acids. Previous experiments involving addition of fatty acid standards to the cell cultures demonstrated that the washing procedure removes >98% of fatty acids not incorporated by the cells (2). Total cellular lipids were extracted from cells or from broken cell extracts by the method of Bligh and Dyer (12) as previously described (2). Phospholipids were fractionated by silicic acid chromatography as previously described (13). Transmethylation of phospholipid fatty acids was done by the method of Morrison and Smith (14). Gas liquid chromatography was performed by on column injection of a Supelcowax 10 capillary column (0.75 mm, inner diameter, × 30 m) using a using a Varian 3700 gas chromatograph at 190 °C using helium as a carrier gas.

*Growth Tests*—Strain L8-14C containing the disrupted *ole1::LEU2* gene was grown overnight at 30 °C with rotary shaking on medium containing 1 mM 18:1 Δ-9. Cells were harvested by centrifugation, washed 3 times in ice-cold distilled water, and inoculated into 20 ml of SDT medium containing the appropriate 1 mM fatty acid supplement at a density of $3 \times 10^5$ cells/ml. Cell density was monitored by hemocytometer counts at 12 and 27 h. If clumping was observed, 1-ml aliquots of the culture were pelleted in a microcentrifuge and washed 2 times with distilled water to disperse the cells before counting.

## RESULTS

Two *OLE1* promoter-β-galactosidase gene fusions were constructed to test the regulation of *OLE1* (Fig. 1). Recombinant plasmid p40 contains a 935-bp *Hin*dIII/*Sal*I fragment of the *OLE1* promoter and 27 N-terminal codons of the protein coding sequence fused in frame to the *E. coli lacZ* gene in expression vector YEp356R. That multiple copy plasmid yields high levels of reporter gene expression and was used for initial experiments to screen fatty acids. Plasmid pCT:OLE contains a *Hin*dIII/*Hpa* fragment of the *OLE1* promoter that includes the gene activation and unsaturated fatty acid regulation sequences but does not contain the *OLE1* "TATA" sequences or the transcription initiation site. That fragment was fused to the single copy CEN plasmid pCT (8) which contains the yeast *CYC1* TATA elements fused to β-galactosidase.

Experiments were performed on cells that were initially in one of two regulatory states. Test fatty acids were added either to cultures in which the reporter gene was derepressed (by growing cells initially without unsaturated fatty acids) or to cultures in which the reporter gene was repressed (cells grown initially in the presence of unsaturated fatty acids). Experiments on initially derepressed cells revealed that relatively long exposure times to fatty acids were required before reporter enzyme activity fell to levels representing the fully repressed state. The second method monitored the relatively rapid synthesis of the reporter enzyme during recovery from the repressed state and yielded larger ratios of derepressed over repressed activities. Cell densities must be maintained at low levels in those experiments, however, to avoid depleting the medium of unsaturated fatty acids while initially repressing the gene.

*Regulation of lacZ Fusion Plasmids by Δ-9 Unsaturated Fatty Acids*—Table I shows the effects of a series of monounsaturated and polyunsaturated fatty acids on reporter gene activity in wild type strain, L8-25A, containing *lacZ* fusion plasmid p40 (Fig. 1*b*). In those experiments cells were grown in selective medium containing oleic acid (18:1 Δ-9) to repress reporter gene activity, washed extensively, and transferred to fresh medium to test for their ability to maintain repression when supplied with different fatty acids.

Comparison of relative enzyme levels revealed that β-galactosidase activity remained repressed in all cultures containing fatty acids with a Δ-9 cis double bond and a hydrocarbon

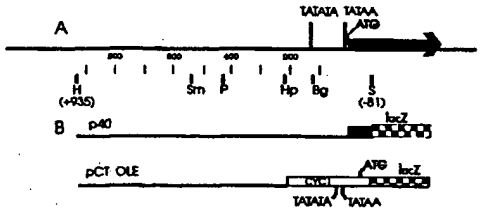



FIG. 1. *A*, restriction map of the *OLE1* upstream and N-terminal coding region. *H, Hind* III; *Sm, Sma*I; *P, Pst*I; *Hp, Hpa*I; *Bg, Bgl*II; *S, Sal*I. *B*, reporter gene constructs containing *OLE1* promoter sequences. Plasmid p40 contains a 935-bp *Hin*dIII/*Sal*I fragment consisting of the *OLE1* promoter region and 81 bp of the N-terminal coding sequence. That fragment was ligated in frame to the *lacZ* coding sequence of multiple copy plasmid YEP356R that also contains the yeast URA3 gene and yeast 2-μm circle ARS sequences. Plasmid pCTOLE was constructed by ligating the *Hin*dIII/*Hpa* fragment containing upstream regions of the *OLE1* promoter but lacking the TATA elements and *OLE1* N-terminal sequences. That fragment was inserted upstream of the TATA elements of the *CYC1-lacZ* fusion vector pCT, a CEN plasmid that contains the *URA3* gene and an ARS1 element.

TABLE I

*Effects of Δ-9 double bond containing monounsaturated and polyunsaturated fatty acids on reporter gene activity from cultures repressed with 18:1*

Strain L8-25A containing plasmid p40 was grown overnight at low density in medium containing 1 mM 18:1; cells were washed and inoculated at 1/10 volume into medium containing 1 mM of the following fatty acids and allowed to grow for an additional 8 h. Cells were again washed and assayed for β-galactosidase activity. β-Galactosidase units are expressed as units/$A_{600}$ (corrected for light scattering).

| Fatty acid supplement | β-Galactosidase activity | No fatty acid control |
|---|---|---|
| | | % |
| None | 38.9 | 100 |
| 18:1 Δ-9 (control)[a] | 0.6 | 1.5 |
| 18:1 Δ-9[b] | 0.8 | 1.2 |
| 16:1 Δ-9 | 0.5 | 1.2 |
| 18:2 Δ-9 | 0.4 | 1.1 |
| 18:3 Δ-6,9,12 | 0.4 | 1.1 |
| 18:3 Δ-9,12,15 | 0.5 | 1.2 |
| 14:1 Δ-9[c] | 0.3 | 0.8 |

[a] Starting culture.
[b] Cells transferred to fresh medium containing 18:1.
[c] Determined in a separate series of experiments.

TABLE II

*Effect of double bond position on OLE1 gene expression*

Strain L8-25A containing the p40 *lacZ* fusion plasmid grown under repressing conditions was tested according to the protocol described in Table I.

| Unsaturated fatty acid supplement | β-Galactosidase activity | No fatty acid control |
|---|---|---|
| | | % |
| 14:1 Δ-9 | 0.068 | 1.8 |
| 18:1 Δ-9 | 1.4 | 3.8 |
| 17:1 Δ-10 | 28.1 | 77 |
| 18:1 Δ-5 | 23.5 | 64 |
| 18:1 Δ-11 | 23.6 | 65.6 |
| No fatty acid | 37.3 | 100 |

chain length from 14–18 carbons. Control cultures containing no fatty acids showed approximately 65-fold greater reporter gene activity. The dienoic species, linoleic acid (18:2, Δ9,12) and two trienoic 18:3 species (Δ-6,9,12 and Δ-9,12,15) were equally strong repressors of the β-galactosidase activity in spite of differences in the positions of the second and third double bonds.

*Regulation of lacZ Fusion Plasmids by Unsaturated Fatty Acids without Δ-9 Double Bonds*—All of the fatty acids tested in the above experiments have a double bond at the Δ-9 position. To test whether a Δ-9 double bond is specifically required for repression, cells were grown under oleic acid repressed conditions and then resuspended in medium that contained fatty acids with double bonds in other positions. Cells grown in medium containing the monounsaturates 17:1 Δ-10, 18:1 Δ-5, or 18:1 Δ-11 had high levels of β-galactosidase activity, similar to those found in control derepressed cultures (Table II). Those activities were more than 16-fold higher than that observed in cells exposed to the repressing fatty acid 18:1 Δ-9.

*Unsaturated Fatty Acid Regulation of an OLE1:CYC1 lacZ Gene Fusion*—The *ole1:lacZ* gene fusion used in the above experiments included elements of *OLE1* mRNA leader sequences and 81 base pairs of the N-terminal coding sequences. To exclude the possibility that translation or mRNA stability might be also controlled through determinants on those parts of the reporter gene message, the effect of double bond position on the regulation of the single copy *lacZ* fusion plasmid pCT:OLE was tested. The reporter gene in that plasmid is under the control of *OLE1* upstream promoter sequences but lacks any *OLE1* mRNA coding sequences (Fig. 1*b*). Cells were grown at low densities under repressed conditions prior to inoculation in various fatty acid containing media (Table III). These data show that 18:2, which contains a Δ-9 bond, strongly represses the reporter gene. Cells exposed to 17:1 Δ-10, 19:1 Δ-10, and 18:1 Δ-11, however, were all derepressed and exhibited β-galactosidase activity comparable with or greater than that of the controls grown in the absence of fatty acids. Thus these data support the results obtained with the p40 plasmid and indicate that transcriptional control accounts for almost all of the regulation of its reporter gene activity.

*Correlation of Fatty Acid Regulation with Growth*—Experiments were carried out to determine if the ability of a fatty acid to act as a repressor is related to its ability to sustain growth of a strain with a disrupted *OLE1* gene (Table IV). Cells were grown for a fixed period on each unsaturated fatty acid and then counted to determine their relative cell densities. The relative densities are compared with the control culture that contained the native 18:1 Δ-9 species which underwent approximately 6.7 doublings during the 27-h experiment. All 16–18 carbon fatty acid species tested that repress the reporter gene were found to support growth of the

TABLE III

*Fatty acid regulation of hybrid OLE1:CYC1:lacZ gene fusion*

Strain L8-25A containing plasmid pCT:OLE was grown under repressed conditions as described in Table I and tested for β-galactosidase activity after transfer and growth for 8 h in medium containing the following supplements. Cells were disrupted and tested for β-galactosidase activity according to the protocol described under "Materials and Methods." Enzyme activity is expressed in units/microgram protein. NFA, no fatty acid.

| Supplement | β-Galactosidase | NFA control |
|---|---|---|
| | | % |
| No fatty acid | 4.32 | 100 |
| 18:2 Δ-9,12 | 0.465 | 10.7 |
| 17:1 Δ-10 | 5.7 | 132 |
| 19:1 Δ-10 | 4.8 | 111 |
| 18:1 Δ-11 | 4.6 | 106 |

TABLE IV

*Effect of fatty acid supplements on growth of ole1::Leu2 disrupted strain L8-14C*

Cells from gene-disrupted *ole1::LEU2* strain L8-14C grown overnight on medium containing 1 mM 18:1 were washed as described under "Materials and Methods" and inoculated into medium containing the listed 1 mM fatty acid supplements at an initial cell density of $5 \times 10^4$ cells/ml.

| Unsaturated fatty acid supplement | 18:1 Δ-9 cell density (12 h) | 18:1 Δ-9 cell density (27 h) |
|---|---|---|
| | % | % |
| None | 2.2 | 0.3 |
| 18:1 Δ-9 | 100 | 100 |
| 16:1 Δ-9 | 98 | 95 |
| 14:1 Δ-9 | 47 | 67.6 |
| 18:2 Δ-9,12 | 121 | 85 |
| 18:3 Δ-6,9,12 | 92 | 85 |
| 18:3 Δ-9,12,15 | 78 | 84 |
| 17:1 Δ-10 | 128 | 62.5 |
| 19:1 Δ-10 | 2 | 3.6[a] |
| 18:1 Δ-5 | 45 | 16.8 |
| 18:1 Δ-11 | 62 | 41.4 |
| 20:1 Δ-11 | 3.5 | 2.2 |
| 20:1 Δ-13 | 1.9 | 0.5 |

[a] No budding cells at 24 h; cultures failed to grow further when incubation was continued for 72 h.

*ole1::LEU2* gene-disrupted strain at levels comparable with those observed when 18:1 Δ-9 was the supplement. 14:1 Δ-9 promoted growth leading to about half the density of the other species (approximately 6 generations). Fatty acids that had no effect or minimal effects on regulation, such as 17:1 Δ-10 and 18:1 Δ-11, also sustained growth at levels similar to that found with medium containing the 14:1 Δ-9 species. Petroselenic acid (18:1 Δ-5) promoted growth at a slightly lower rate than the Δ-9-containing species. The longer chain, non-repressing fatty acids 19:1 Δ-10, 20:1 Δ-11, and 20:1 Δ-13 did not support significant growth in tests that extended for 72–96 h.

*Incorporation of Exogenous Fatty Acids into Membrane Lipids of OLE1⁺ Cells*—Since transcriptionally non-repressing fatty acids such as 17:1 Δ-10 and 18:1 Δ-11 can repair the growth requirements of the *ole1* gene disrupted strain they must enter the cells and be incorporated into membrane lipids. Their inability to repress *OLE1* transcription could be due to the fact that those fatty acids are selectively excluded from strains containing a functional desaturase or that they are modified to another species upon entry into the cell. To examine that possibility, *OLE1* cultures were analyzed to determine if non-repressing fatty acids were incorporated into cellular total lipid and phospholipid fractions. Cultures of L8–25A (relevant genotype *OLE1*) were grown under derepressed conditions overnight and then transferred into minimal medium or medium containing those fatty acids for 12 h. Gas chromatograms of total lipid fatty acids extracted from washed cells are shown in Fig. 2, and their relative levels are shown in Table V.

Cells grown on minimal medium (Fig. 2*a*) contained the normal distribution of fatty acids in which 16:1 and 18:1 comprise greater than 75% of the total fatty acids. Fig. 2*b* indicates the fatty acid distribution when 17:1 Δ-10 is added to the medium. Surprisingly, 17:1 Δ-10 is the dominant species, accounting for almost 80% of the total cellular fatty acids. Since that fatty acid does not have a significant effect on reporter gene transcription we expected the desaturase to be active at normal, derepressed levels. This would be indicated by high levels of the 16:1 and 18:1 enzyme products in the lipid fraction. Those fatty acid levels were in fact strongly depressed (compare their abundance with 16:0 and 18:0), indicating that in the presence of 17:1 Δ-10, desaturase enzyme activity is sharply reduced. In separate experiments the same distribution of fatty acids was found in the phospholipid fraction of 17:1 Δ-10 fed cells, as determined by analysis of lipid fractions separated by silicic acid chromatography. 17:1 Δ-10 comprised approximately 72% of phospholipid fatty acids in that test, while 16:1 and 18:1 represented approximately 5.5% of the total (data not shown).

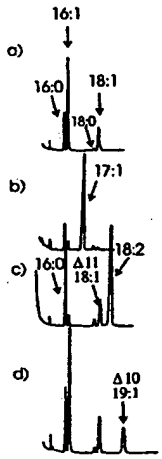



FIG. 2. **Gas chromatograms of total lipid fatty acids from strain L8–25A grown at 30 °C with repressing and non-repressing fatty acids.** Cells were inoculated at low density ($5 \times 10^4$/ml) in 100 ml of medium containing either no fatty acids (*a*) or 1 mM fatty acid supplements (*b–d*). Cultures were harvested at a density of $1 \times 10^7$/ml, and the extensively washed cell pellets were subjected to HCl methanolysis as described under "Experimental Procedures." Fatty acid supplements: *a*, no fatty acid control; *b*, 1 mM 17:1, Δ-10; *c*, 0.25 mM 18:1, Δ-11 + 0.75 mM 18:2, Δ-9,12; *d*, 1 mM 19:1,Δ-10.

TABLE V

*Fatty acid composition of L8-25A grown on non-repressing fatty acids*

| Culture supplement | 16:0 | 16:1, Δ-9 | 18:0 | 18:1, Δ-9 | Fed fatty acid | Total UFA[a] |
|---|---|---|---|---|---|---|
| No fatty acid | 20.1 | 50.3 | 2.3 | 22.4 | | 77.7 |
| 17:1 Δ-10 | 12.1 | 4.3 | 2.8 | 1.8 | 79.0 | 85.1 |
| 18:1 Δ-11 | 24.2 | 25.6 | 4.4 | | 40.5[b] | 66.1 |
| 18:1 Δ-11 + 18:2 Δ-9,12 | 20.3 | 2.7 | 3.3 | Trace | 11.9[c]<br>60.5[d] | 75.1 |
| 19:1 Δ-10 | 17.3 | 44.1 | 2.7 | 17.3 | 16.1 | 77.5 |

[a] UFA, unsaturated fatty acid.
[b] 18:1 Δ-11 and 18:1 Δ-9 in approximately equimolar amounts.
[c] 18:1, Δ-11 and a small fraction of 18:1 Δ-9; see text for rationale and procedures.
[d] 18:2 Δ9,12.

Analysis of cells fed 18:1 Δ-11 (Table V) indicated that it was incorporated at moderate levels in cellular lipids. Unlike the effects seen with 17:1 Δ-10 fed cultures, however, cellular 16:1 and 18:1 levels were not severely reduced (indicating that the desaturase activity was not repressed in the presence of that fatty acid). Due to overlapping peaks of the endogenous 18:1 Δ-9 and the fed 18:1 Δ-11, it was not possible to accurately quantify their relative levels in lipids from cultures fed only the Δ-11 species. To further examine the incorporation of the 18:1 Δ-11, a second series of experiments was done in which cells were fed a mixture of 18:2 Δ-9,12 and 18:1 Δ-11 (Fig. 2*c*). The 18:2 should repress the production of endogenous 16:1 and 18:1 formed by the desaturase, thus revealing the presence of the incorporated Δ-11 species (Fig. 2*c*). An approximate 3:1 mixture of the 18:2 and 18:1 species was added to the media and analysis of the lipids revealed that the two fed species comprised approximately 70% of the total fatty acids in that ratio. The retention time of the 18:1 peak also corresponds to that of the Δ-11 species although a shoulder on the chromatogram suggests that a small amount of endogenous 18:1 may also be present. The sharp loss of 16:1 clearly indicates that the desaturase activity was strongly repressed under those growth conditions.

Although 19:1 Δ-10 did not repress transcription of the *OLE1* reporter gene and was incapable of supporting growth of the disruption strain, that fatty acid was readily incorporated into cellular lipids (Fig. 2*d*) and comprised about 16% of the total fatty acids. Endogenous 16:1 and 18:1 were slightly reduced under those conditions, resulting in total cellular unsaturated fatty acid levels that were similar to those found in cells grown with no fatty acids in the culture medium.

*Effect of Fatty Acids on OLE1 mRNA Levels*—*OLE1* mRNA levels were measured to determine the response to repressing and non-repressing fatty acids. In those experiments cells were grown under derepressed conditions to mid-logarithmic phase, shifted to fatty acid containing medium, and allowed to grow for an additional 4 h before harvesting and RNA isolation. Previous studies had shown that *OLE1* mRNA is reduced to low levels within 15 min under repressing conditions (1). *OLE1* mRNA levels were compared with ribosomal

subunit L32 mRNA as an internal control for loading.

Fig. 3 shows that *OLE1* transcript levels in cultures grown in 18:1 Δ-11 were similar to those in control cultures containing no fatty acids. Cultures grown under the same conditions in the presence of 18:2 Δ-9,12 and 17:1 Δ-10, however, contained sharply reduced *OLE1* mRNAs that were approximately 1/20 the levels found under derepressed conditions, indicating that both species strongly repress available *OLE1* message.

## DISCUSSION

The studies of *OLE1* gene expression reported in this paper are designed to dissect the mechanisms of unsaturated fatty acid regulation in yeast. A variety of unsaturated fatty acids have been fed to yeast cells to determine and correlate their individual effects on: (*a*) *OLE1* native mRNA levels, (*b*) *OLE1* transcription *per se* (as measured by the activity of a reporter gene under the control of the *OLE1* promoter, and (*c*) the relative activity of the Δ-9 desaturase enzyme as determined by the level of its products (16:1 and 18:1) in cellular lipids.

We have previously reported that the addition of unsaturated fatty acids to the growth medium strongly represses Δ-9 desaturase (*OLE1*) mRNA levels. The data presented here show that one component of that regulatory response occurs at the level of transcription since β-galactosidase gene fusions containing only *OLE1* upstream promoter sequences are strongly repressed by certain unsaturated fatty acids. Although the transcriptional regulation is triggered by a range of fatty acids that vary with respect to chain length and number of double bonds, it has a highly specific requirement for a double bond in the Δ-9 position. This suggests that this regulatory mode is initiated by the binding of a fatty acid to a protein sensor rather than as a response to changes in the physical characteristics of the membrane lipid bilayer caused by the increased availability of unsaturated fatty acids. The sensor protein must apparently recognize the carboxyl (or carbonyl group) of the hydrocarbon chain and the Δ-9 double bond, since it is insensitive to chain length and additional double bonds distal to that position. It appears to tolerate some differences in the structure of the chain between carbons 1 and 9, however, since 18:3 Δ-6,9,12 is a highly effective repressor in spite of the presence of a double bond at the Δ-6 position.

This specificity of the transcriptional sensor was surprising in light of previous studies (4) which demonstrated that a much wider range of Δ-9 and non-Δ-9 double bond-containing unsaturated fatty acids satisfy the growth requirements of *OLE1* mutant strains. Data presented here suggest that the transcriptional regulation of *OLE1* is one component of a more complex system that regulates desaturase activity and controls the composition of unsaturated fatty acids in membrane lipids. This is particularly evident from experiments that show that 17:1 Δ-10, which does not regulate transcription, strongly represses *OLE1* mRNA levels and desaturase activity. The absence of *OLE1* message in those cells in the presence of continued transcriptional activity points to the existence of a post-transcriptional mechanism that may control mRNA stability. The identification of 17:1 Δ-10 as a specific stimulus for post-transcriptional regulation of the *OLE1* gene will be useful in defining the characteristics of that system.

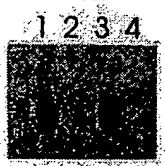



FIG. 3. RNA blot hybridization of total RNA from strain L8-25A grown to mid-logarithmic phase on minimal medium and then exposed for 4 h to 1 mM unsaturated fatty acids. 100 μg of total cellular RNA was loaded for each fraction. *1*, control (no fatty acids); *2*, 18:2 Δ-9,12; *3*, 18:1 Δ-11; *4*, 17:1 Δ-10. Blots were probed with the entire protein coding sequences of the *OLE1* gene.

Additional controls may also exist that affect the activity or the stability of the desaturase enzyme. These are suggested by the pattern of fatty acid integration in wild type and *OLE1* mutant cultures that were fed 18:1 Δ-11. In the wild type cells *OLE1* transcription and mRNA levels are unaffected by that species, yet the integration of the non-repressing fatty acid into membrane lipids results in a reduction of endogenous 16:1 and 18:1 levels suggesting that the activity of the enzyme is modulated by the presence of the exogenously supplied acid. This ability of cells to maintain balanced ratios of saturated and unsaturated fatty acids under widely differing conditions give further evidence that there are finely tuned controls that regulate membrane fatty acid composition. It is somewhat surprising that 17:1 Δ-10 triggers a strong post-transcriptional repression of the gene, whereas 18:1 Δ-11 and 19:1 Δ-10 do not. This may indicate that 17:1 Δ-10 repression is triggered by the physical characteristics of membrane lipids containing that species rather than by the specific recognition of the fatty acid. The intermediate chain length of the 17-carbon species may mimic the normal distribution of 16:1 and 18:1 in maintaining appropriate membrane fluid properties whereas the longer chain fatty acids may have a rigidifying effect on the phospholipid bilayer, triggering the need for additional 16:1 and 18:1. The regulatory sensor that detects the properties of a "normal" membrane would then be responsible for initiating the destabilization of the *OLE1* message.

The question arises as to where in the cell the sensors for these regulatory systems reside and what form of fatty acid triggers each response. Sensory elements that are associated with the endoplasmic reticulum, for example, could be expected to be part of a system that regulates the composition and fluid properties of the lipid bilayer, while elements that are soluble cytoplasmic proteins might represent a simpler form of metabolic control. Since exogenous fatty acids transported into the cell are presumably converted to CoA derivatives, it is possible that they could be the regulatory stimulus for a cytoplasmic sensor. Transport of unsaturated fatty acids into the cell would presumably increase the proportion of unsaturated species in the long chain acyl-CoA pool, providing the regulatory stimulus. In the simplest case the transcriptional regulatory sensor might be a domain of a soluble unsaturated acyl-CoA-binding protein that can be transported to the nucleus and act directly on the *OLE1* transcription apparatus. This is analogous to transcriptional regulators that are responsive to steroid hormones such as the glucocorticoid receptor (15, 16). Fatty acid-binding proteins have been recently identified in nearly all mammalian tissues and are generally found as abundant cytosolic proteins (17). Studies of the binding affinities of the two most well characterized proteins I-FABP and L-FABP indicate that they bind a wide degree of saturated and unsaturated fatty acids as well as other hydrophobic ligands (18). This suggests that they lack the necessary specificity found with the *OLE1* transcriptional sensor but does not rule out the existence of other acyl-binding proteins that have a high specificity for an unsaturated species.

It is equally plausible to suggest that sensors for transcriptional and/or post-transcriptional mechanisms may act at the level of the membrane and recognize fatty acids that are

acylated to glycerolipids. These might function to maintain the balance of saturated and unsaturated fatty acids in membrane lipids. If the transcriptional regulatory sensor identified here is an integral membrane protein, other proteins would be required to complete the regulatory circuit. Post-transcriptional sensors, however, could act locally at the endoplasmic reticulum to regulate the translation or stability of the *OLE1* message or by modulation of enzyme activity.

The regulation of the Δ-9 enzyme in yeast has similarities to its regulation in liver and adipocytes (21–23). Liver desaturase mRNA levels have been shown to be regulated in response to a variety of dietary lipids. In fact, a wide variety of organisms ranging from microbes to mammalian cells incorporate exogenous fatty acids into cellular lipids (23, 24), and specific mechanisms have evolved to transport these molecules across the plasma membrane (19, 20). This appears to be part of a system generally used by cells to bypass the energetically expensive synthesis of fatty acids (which make up a large part of the cellular mass) by preferentially importing saturated and unsaturated fatty acids from the growth medium. Evidence presented in this paper indicates that the regulation of fatty acid desaturation involves a complex circuit that balances external and internal fatty acids utilization with the physiological requirements of the cell. Another important function of this regulation appears to be involved in the maintenance of glycerolipid fatty acyl composition.

The existence of multiple sensors and regulatory paths for the *OLE1* gene that exert a graduated degree of control on desaturase activity appear to parallel other membrane lipid biosynthetic enzymes. Similar multiple levels of control have been described for 3-hydroxy-3-methylglutaryl-coenzyme A reductase, and these have been cited as evidence of its key regulatory role in sterol metabolism (25). Dissection of the components of these regulatory circuits should yield information concerning the mechanisms that balance the synthesis of membrane lipid species.

*Acknowledgments*—We gratefully acknowledge P. Flanagan and R. Kornberg for generously providing us with plasmid pCT and J. Warner for the plasmid containing the gene L32.

## REFERENCES

1. Bossie, M. A., and Martin, C. E. (1989) *J. Bacteriol.* **171**, 6409–6413
2. Stukey, J. E., McDonough, V. M., and Martin, C. E. (1989) *J. Biol. Chem.* **264**, 16537–16544
3. Stukey, J. E., McDonough, V. M., and Martin, C. E. (1990) *J. Biol. Chem.* **265**, 20144–20149
4. Henry, S. A. (1982) in *The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression* (Strathern, J. N., Jones, E. W., and Broach, J. R., eds) pp. 101–158, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
5. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1987) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York
6. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, pp. 97–164, Cold Spring Harbor, Cold Spring Harbor, NY
7. Eng, F. J., and Warner, J. R. (1991) *Cell* **65**, 797–804
8. Lue, N. F., and Kornberg, R. D. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 8839–8843
9. Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254
10. Browse, J., McCourt, P., and Somerville, C. R. (1986) *Plant Physiol.* **81**, 859–864
11. Kates, M. (1972) *Techniques of Lipidology*, pp. 127–128, American Elsevier, New York
12. Bligh, E. G., and Dyer, W. J. (1959) *Can. J. Biochem. Biophys.* **37**, 911–917
13. Martin, C. E., Siegel, D., and Aaronson, L. R. (1981) *Biochem. Biophys. Acta* **665**, 399–407
14. Morrison, W. R., and Smith, L. M. (1964) *J. Lipids Res.* **5**, 600–608
15. Beato, M. (1991) *FASEB J.* **5**, 2044–2051
16. Godowski, P. J., and Picard, D. (1989) *Biochem. Pharmacol.* **38**, 3135–3143
17. Clarke, S. D., and Armstrong, M. K. (1989) *FASEB J.* **3**, 2480–2487
18. Nemecz, G., and Schroeder, F. (1991) *J. Biol. Chem.* **266**, 17180–17186
19. Nunn, W. D. (1986) *Microbiol. Rev.* **50**, 179–192
20. Storch, J., Lechene, C., and Kleinfeld, A. M. (1991) *J. Biol. Chem.* **266**, 13473–13476
21. Nakatani, Y., Horishoki, M., Brenner, M., Yamamoto, T., Besnard, F., Roeder, R. G., and Freese, E. (1990) *Nature* **348**, 86–88
22. Brenner, R. R. (1989) in *The Role of Fats in Human Nutrition* (Vergroesen, A. J., and Crawford, M., eds) 2nd Ed., pp. 45–79, Academic Press, New York
23. Thiede, M. A., and Strittmatter, P. (1985) *J. Biol. Chem.* **260**, 14459–14463
24. Strittmatter, P., Thiede, M. A., Hackett, C. S., and Ozols, J. (1988) *J. Biol. Chem.* **263**, 2532–2535
25. Goldstein, J. L., and Brown, M. S. (1990) *Nature* **343**, 425–430

# EXHIBIT B

*Proc. Natl. Acad. Sci. USA*
Vol. 90, pp. 2486–2490, March 1993
Biochemistry

# Stearoyl-acyl carrier protein $\Delta^9$ desaturase from *Ricinus communis* is a diiron-oxo protein

(fatty acid desaturation/binuclear iron cluster/Mössbauer spectroscopy/iron-binding motif)

BRIAN G. FOX*†, JOHN SHANKLIN†‡§, CHRIS SOMERVILLE‡, AND ECKARD MÜNCK*

*Department of Chemistry, Carnegie Mellon University, Pittsburgh, PA 15213; and ‡Department of Energy Plant Research Laboratory, Michigan State University, East Lansing, MI 48824

*Communicated by Helmut Beinert, November 30, 1992*

ABSTRACT **A gene encoding stearoyl-acyl carrier protein $\Delta^9$ desaturase (EC 1.14.99.6) from castor was expressed in *Escherichia coli*. The purified catalytically active enzyme contained four atoms of iron per homodimer. The desaturase was studied in two oxidation states with Mössbauer spectroscopy in applied fields up to 6.0 T. These studies show conclusively that the oxidized enzyme contains two (identical) clusters consisting of a pair of antiferromagnetically coupled ($J > 60$ cm$^{-1}$, $H = JS_1 \cdot S_2$) $Fe^{3+}$ sites. The diferric cluster exhibited absorption bands from 300 to 355 nm; addition of azide elicited a charge transfer band at 450 nm. In the presence of dithionite, the clusters were reduced to the diferrous state. Addition of stearoyl-CoA and $O_2$ returned the clusters to the diferric state. These properties are consistent with assigning the desaturase to the class of $O_2$-activating proteins containing diiron-oxo clusters, most notably ribonucleotide reductase and methane monooxygenase hydroxylase. Comparison of the primary structures for these three catalytically diverse proteins revealed a conserved pair of the amino acid sequence -(Asp/Glu)-Glu-Xaa-Arg-His- separated by ≈100 amino acids. Since each of these proteins can catalyze $O_2$-dependent cleavage of unactivated C—H bonds, we propose that these amino acid sequences represent a biological motif used for the creation of reactive catalytic intermediates. Thus, eukaryotic fatty acid desaturation may proceed via enzymatic generation of a high-valent iron-oxo species derived from the diiron cluster.**

Fatty acids are synthesized in the plastid of higher plants by the acyl carrier protein (ACP) pathway (1). The penultimate product, stearoyl-ACP, is desaturated to oleoyl-ACP by the soluble enzyme stearoyl-ACP $\Delta^9$ desaturase (EC 1.14.99.6) in the presence of $O_2$, NAD(P)H, NAD(P)H ferredoxin oxidoreductase, and ferredoxin (2). All other known desaturases are integral membrane proteins which act upon membrane lipids. The soluble desaturase has been purified from several plant species (3–5) and shown to be a homodimer of ≈70 kDa (5). Comparison of the amino acid sequences deduced from cDNA clones (3, 4) revealed that the stearoyl $\Delta^9$ desaturase from distantly related plant species is a highly conserved polypeptide. In contrast, no substantial homology was observed with the corresponding enzymes from animals (6).

Previous studies have shown that the desaturase (7) contains iron and is inhibited by iron chelators and cyanide, but not by carbon monoxide (2, 7). In addition, both forms of stearoyl $\Delta^9$ desaturase (5, 7) exhibit absorbance features between 300 and 400 nm and no Soret absorption. However, detailed analysis of the metal content and the optical features have not yet been reported. Here we detail the engineering of the castor (*Ricinus communis*) desaturase into a bacterial expression system which provides large quantities of functional enzyme. Protein and metal content determinations demonstrate that the desaturase contains four atoms of catalytically essential iron. Through the use of optical and Mössbauer (see ref. 8 for a general introduction) spectroscopies, this iron is shown to reside in a diiron-oxo cluster with properties similar to clusters of hemerythrin, ribonucleotide reductase, rubrerythrin, purple acid phosphatase, and the methane monooxygenase (MMO) hydroxylase (see ref. 9 for a review).

The MMO hydroxylase (10), ribonucleotide reductase (11), and the yeast stearoyl-CoA $\Delta^9$ desaturase (12) catalyze oxygenase reactions, albeit with distinctly different chemical outcomes. For these three oxygenases, high-valent iron-oxo structures have been proposed as catalytic intermediates. A comparison of the primary structures of the castor desaturase (3), MMO hydroxylase (13), and ribonucleotide reductase (14) presented here has revealed the presence of a highly conserved set of carboxylate and histidine ligands which may constitute the iron-containing active site. Consequently, oxidative desaturation may also involve generation of a high-valent iron-oxo species derived from the diiron-oxo cluster.

## MATERIALS AND METHODS

**Plasmid Construction.** An open reading frame corresponding to the mature castor desaturase was identified by comparison of the deduced amino acid sequences of the castor and safflower desaturases. A region of the castor cDNA (GenBank accession no. M59857) was amplified by PCR using oligonucleotides 5′-TTAACCATGGCCTCTACCCTCAAG and 3′-CGATCCATGGATCGTTTGCTTATTA. *Nco* I restriction sites were introduced at the 5′ (corresponding to Met-32) and 3′ (base pair 1565) ends of the amplified region, such that the resulting sequence could be digested and inserted into the *Nco* I site of the expression vector pET3d. The resulting plasmid (pRCMD9) was used to transform the *Escherichia coli* strain BL21(DE3).

**Growth and Expression.** The medium used for bacterial growth contained (per liter): Bacto tryptone, 10 g; NaCl, 5 g; and ampicillin, 50 mg. For Mössbauer studies, $^{57}Fe$ metal was dissolved in a minimal volume of 1 M HCl and added to provide ≈80% isotopic enrichment. Cells were grown to $OD_{600} = 0.5$ at 37°C. Expression was induced by addition of isopropyl β-D-thiogalactopyranoside (0.4 mM). Four hours after induction, cells were collected by centrifugation, washed in 40 mM Tris·HCl, pH 8.0, and stored at −70°C.

**Purification.** Cell paste (≈20 g) suspended in 50 ml of 40 mM Tris·HCl, pH 8.0, buffer was disrupted by a French pressure cell. The disrupted cell suspension was diluted 4-fold with buffer and DNase I (5 mg) was added. This

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: ACP, acyl carrier protein; MMO, methane monooxygenase.

†To whom reprint requests should be addressed.
§Present address: Department of Biology, Brookhaven National Laboratory, Upton, NY 11973.

mixture was centrifuged at 37,000 × *g* for 45 min at 4°C. The supernatant was diluted 2-fold with buffer and applied to a DEAE-Sephacel column (50 × 150 mm; Pharmacia) at a rate of 30 cm/hr. After loading, the column was washed with 1 liter of buffer. A 1-liter gradient of 0.0–0.3 M NaCl in 40 mM Tris·HCl, pH 8.0, buffer was applied at a rate of 15 cm/hr. The desaturase was eluted at ≈0.1 M NaCl. Fractions containing the desaturase were identified by SDS/PAGE and concentrated by ultrafiltration. The concentrated protein was applied to a Sephadex G-75 column (25 × 1000 mm) equilibrated with 50 mM Hepes, pH 7.8, containing 50 mM NaCl and 5% (vol/vol) glycerol and eluted at a rate of 5 cm/hr. Desaturase fractions were identified and concentrated as before and stored at −70°C.

**Activity Assay.** Activity was measured as the ferredoxin-dependent release of $^3H_2O$ as described in ref. 3, except that [9,10-$^3$H]stearoyl-ACP was used in place of [9,10-$^3$H]-stearoyl-CoA.

**Amino Acid and Metal Content.** Extinction coefficients were determined by combined optical spectroscopy and amino acid analysis of purified desaturase protein ($\varepsilon_{278}$ = 80,410 $M^{-1}\cdot cm^{-1}$ and $\varepsilon_{340}$ = 8400 $M^{-1}\cdot cm^{-1}$). Samples of known protein concentration were also subjected to metal analysis by atomic absorption spectroscopy. For correlation of iron content with catalytic activity, a desaturase sample was treated with *o*-bathophenanthroline disulfonate (5 mM). At appropriate time intervals (0.25–16 hr), iron removed from the desaturase was determined optically at 520 nm. The chelator-treated desaturase was then desalted by using a Sephadex G-50 spun column. The protein concentration of the chelator-free desaturase was determined optically and the remaining desaturase activity was determined as described above.

**Spectroscopic Methods.** Concentrations of the desaturase were determined by optical spectroscopy. Sodium azide (4 M in 100 mM Hepes, pH 7.8) was slowly added to the desaturase to a final concentration of 0.8 M. The desaturase was reduced by the addition of dithionite solution (≈1 electron per iron) followed by equilibration for 15 min at 25°C.

## RESULTS

**Physical and Catalytic Properties.** A recombinant castor stearoyl-ACP desaturase lacking the putative transit peptide was constructed, using Met-32 as the translation initiation site. This construct resulted in the addition of one amino acid to the mature desaturase (3). Expression of this cDNA allowed the accumulation of the desaturase to 15–30% of total soluble protein. Following a two-step purification, ≈250 mg of desaturase was obtained from 20 g of cell paste. As judged by SDS/PAGE, the desaturase was greater than 95% pure and appeared as a single band of 37 kDa. Gel filtration measurements showed the recombinant desaturase had molecular mass of 75 kDa, consistent with the $\alpha_2$ quaternary structure observed for all other higher plant stearoyl-ACP $\Delta^9$ desaturases. The catalytic activity (≈10 milliunits/mg) was comparable to that of the best previously available enzyme preparations. These characterizations demonstrate that the recombinant plant desaturase has been correctly assembled into a catalytically active homodimer fully representative of the native plant enzyme.

**Amino Acid and Metal Content.** The measured amino acid content closely matched that predicted from the cDNA sequence, indicating a molecular mass of 83,550 Da. Correlated amino acid and metal analysis showed that the desaturase contained 3.85 ± 0.25 mol of iron per mol of holoprotein. No other metals were detected by atomic absorption, and no flavin, inorganic sulfide, or heme was present. Previous studies have suggested the requirement for iron in the soluble desaturase reaction (2). However, both the ferredoxin reductase and the ferredoxin contain cofactors or metal ions which may potentially interact with these inhibitors. As shown in Fig. 1, the progressive removal of iron from the purified desaturase resulted in the concomitant loss of enzymatic activity. Thus, we conclude iron is essential for stearoyl-ACP $\Delta^9$ desaturase activity. At present, reincorporation of iron into the apoprotein has not been achieved.

**Spectroscopic Characterization.** Absorbance features (solid line of Fig. 2) were observed at 300–355 nm and 475 nm. Addition of sodium azide gave rise to a new complex with an absorption maximum at 345 nm (broken line in Fig. 2). In addition, the weak band observed at 475 nm was replaced by a more intense band at 450 nm. The optical properties described here are essentially the same as those of the corresponding complexes of methemerythrin and ribonucleotide reductase, proteins known to contain $\mu$-oxo-bridged diiron clusters with a primary ligation sphere of oxygen and nitrogen ligands.

Optical spectra (Fig. 3) show that ≈4 electrons per molecule of desaturase were required to complete the reductive titration. The $\varepsilon_{340}$ = 8000 $M^{-1}\cdot cm^{-1}$ calculated from the *Inset* to Fig. 3 is within 5% of that determined by optical spectroscopy and amino acid analysis, suggesting that all iron contained in the desaturase was $Fe^{3+}$ and was converted to $Fe^{2+}$ during the titration. No intermediates with a distinct optical spectrum were observed during the reductive titration. Likewise, no EPR signals were observed when less than 4 electrons were added per molecule of desaturase. The absence of low-temperature EPR signals in the desaturase strongly suggests that all iron must be present in complexes of either integer- or zero-electronic spin.

Fig. 4 shows Mössbauer spectra of the $^{57}$Fe-enriched desaturase recorded at 4.2 K in zero magnetic field. The spectrum of the dithionite-reduced enzyme (Fig. 4, spectrum A) consists of a doublet with (average) quadrupole splitting, $\Delta E_Q \approx 3.2$ mm/s, and isomer shift, $\delta$ = 1.30 mm/s. These parameters are typical of high-spin $Fe^{2+}$ in a 5- or 6-coordinate environment of oxygen and nitrogen ligands. The 180-K data (not shown) indicate that the spectrum consists of two doublets of equal proportion. In fact, the 4.2-K spectrum is best represented as a superposition of two doublets. Least-squares fitting parameters are listed in Table 1. Fig. 4, spectrum B, shows the Mössbauer spectrum of the desaturase as isolated under aerobic conditions. Comparison with spectrum A shows that the as-isolated enzyme contained a fraction (19%) of high-spin $Fe^{2+}$ with the same Mössbauer parameters as the dithionite-reduced desaturase. Examination of the optical spectrum obtained from this sample revealed that the ratio $A_{340}/A_{278}$ was lower than expected (16%) from the extinction coefficient measurements, suggesting the high-spin $Fe^{2+}$ fraction in the as-isolated enzyme was the reduced desaturase. The remaining iron in the as-isolated enzyme was contained in two quadrupole doublets. The majority species, species 1, representing ≈66% of total iron

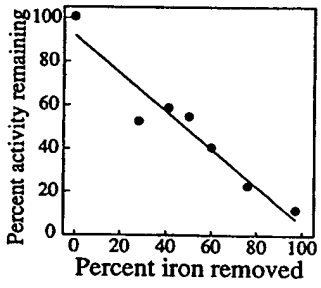



FIG. 1. Correlation between iron content and enzymatic activity. Protein data: 3.8 mol of iron per mol of protein; specific activity of 16 milliunits/mg.

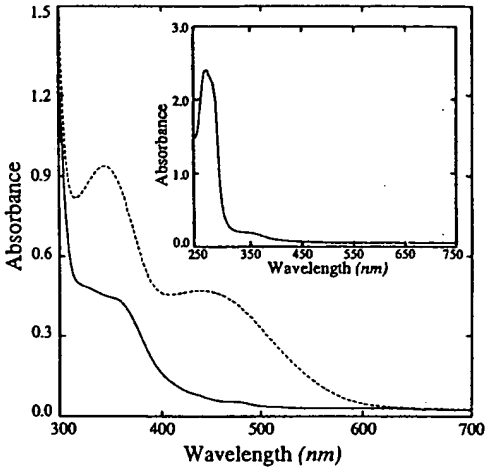



FIG. 2. Optical spectra of the oxidized desaturase (solid line) and the oxidized desaturase plus 0.8 M sodium azide (broken line). Protein data: 50 μM protein; 3.9 mol of iron per mol of protein. (*Inset*) Oxidized desaturase, expanded absorbance scale.

(or 81% of nonferrous iron), has $\Delta E_Q$ = 1.54 mm/s and $\delta$ = 0.53 mm/s, whereas the minority species, species 2, representing ≈15% of total iron, has $\Delta E_Q$ = 0.72 mm/s and $\delta$ = 0.50 mm/s. The splittings of both doublets were found to be independent of temperature in the range 4.2–180 K. The solid line drawn through the data is a least-squares fit to the contributions of the reduced enzyme and species 1 and 2; the contribution of species 2 has been outlined above the data.

Fig. 5 shows 4.2-K Mössbauer spectra of the dithionite-reduced (A) and as-isolated desaturase recorded in a 6.0-T (tesla) applied field. For clarity, the 19% contribution of the reduced enzyme has been subtracted from the raw data of the as-isolated enzyme. The resulting spectrum (B) has features typical of diamagnetic material. This is confirmed by a computer simulation (solid line drawn through spectrum B of Fig. 5) generated with the assumption that the iron atoms of both species 1 and 2 reside in environments with electronic spin $S$ = 0.

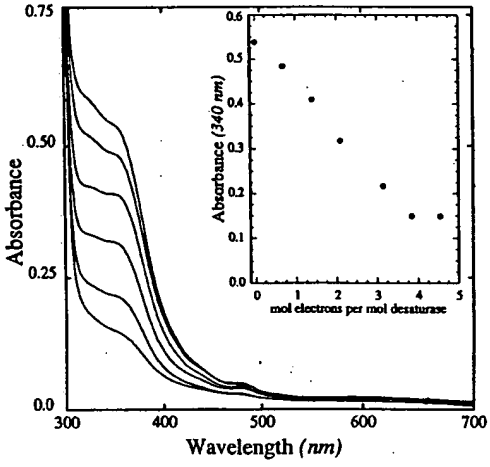



FIG. 3. Reductive titration of the desaturase. (*Inset*) Decrease in optical absorbance at 340 nm versus mol of electrons added per mol of desaturase. The absorbance of the final titration point includes contributions from both the methyl viologen radical state and sodium dithionite anion. Protein data: 79 nmol of protein in 0.85 ml of 50 mM Hepes, pH 7.8, containing 5 μM methyl viologen; 3.9 mol of iron per mol of protein.

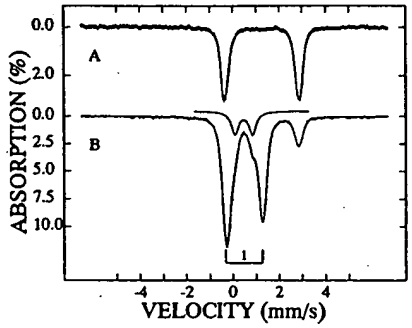



FIG. 4. Zero-field Mössbauer spectra of the dithionite-reduced (A) and as-isolated (B) desaturase recorded at 4.2 K. The solid lines are least-squares fits to the spectra using the parameters listed in Table 1. In B, the doublet of species 1 is indicated by the brackets; the doublet of species 2 is drawn separately above the data. Protein data: 250 nmol of protein in 0.25 ml of 50 mM Hepes, pH 7.8; 3.9 mol of iron per mol of protein.

The Mössbauer properties of both species 1 and 2 are compatible with either a low-spin ($S$ = 0) $Fe^{2+}$ configuration or antiferromagnetically spin-coupled $Fe^{3+}$. A low-spin $Fe^{2+}$ assignment is chemically unreasonable, however, since reduction of the desaturase by ≈1 electron per iron produces the high-spin $Fe^{2+}$ species of Fig. 4, spectrum A. The present observations can be reconciled if species 1 and 2 represent antiferromagnetically coupled high-spin $Fe^{3+}$ sites.[¶] Therefore, we suggest that species 1 and 2 represent the diferric cluster of the oxidized desaturase. The Mössbauer properties of the desaturase shown in Table 1 strongly resemble those of the diiron-oxo clusters found in rubrerythrin, the MMO hydroxylase, ribonucleotide reductase, hemerythrin, and purple acid phosphatase (9).

The oxo-bridge present in the diferric clusters of hemerythrin and ribonucleotide reductase provides a strong antiferromagnetic exchange pathway (exchange coupling constant $J > 200$ cm$^{-1}$, $H_{ex} = JS_1 \cdot S_2$, $S_1 = S_2 = 5/2$) between the two ferric sites. In contrast, the diferric cluster of the MMO hydroxylase (bridging ligands unknown, but probably not oxo) exhibits relatively weak exchange coupling ($J \approx 15$ cm$^{-1}$) (B.G.F. and E.M., unpublished data). As a result of the weaker coupling, the first excited spin multiplet ($S$ = 1) of the diferric MMO hydroxylase is appreciably populated at 20 K. Consequently, Mössbauer spectra taken in strong applied fields exhibit paramagnetic structure associated with the $S$ = 1 multiplet. This property can be used to estimate $J$ by Mössbauer spectroscopy. For the oxidized desaturase, the 6.0-T spectrum recorded at 60 K was identical to spectrum B in Fig. 5, suggesting that $J > 60$ cm$^{-1}$.

The 6.0-T spectrum of the reduced desaturase (Fig. 5, spectrum A) shows that the electronic ground state of the diferrous cluster is paramagnetic, as demonstrated by the presence of sizable magnetic hyperfine interactions (for comparison, a spectrum calculated for a diamagnetic diferrous cluster is shown above the data). The observation of paramagnetic hyperfine structure rules out antiferromagnetic coupling between the ferrous sites that is larger than the

[¶]The values of $\Delta E_Q$ and $\delta$, the temperature independence of $\Delta E_Q$, and the observed diamagnetism all suggest that species 2, like species 1, represents a diferric cluster. If species 2 would represent a low-spin $Fe^{2+}$ species, it should persist after the addition of sodium dithionite, in contrast to the experimental result of Fig. 4, spectrum A. We have recently observed pH-dependent equilibrium changes in the quadrupole patterns of the diferric cluster of the MMO hydroxylase (B.G.F. and E.M., unpublished data). Thus, the presence of species 1 and 2 could represent a pH-dependent equilibrium.

Table 1. Mössbauer properties of the desaturase

| State | $\delta$, mm/s | $\Delta E_Q$, mm/s | $T$, K |
|---|---|---|---|
| Diferrous | 1.24 | 2.75 | 180 |
| | 1.24 | 3.24 | 180 |
| | 1.30 | 3.04 | 4.2 |
| | 1.30 | 3.36 | 4.2 |
| Diferric | | | |
| Species 1 (81%) | 0.53 | 1.54 | 4.2 |
| Species 2 | 0.50 | 0.74 | 4.2 |

Isomer shifts are relative to iron metal at 298 K. Values are for two inequivalent iron sites of equal proportion in the diferrous state.

zero-field splittings of the ferrous ions, suggesting that $J < 15$ $cm^{-1}$ for the diferrous cluster.

Although the high-spin $Fe^{2+}$ component of Fig. 4, spectrum B, has parameters similar to the diferrous cluster, this component could arise from adventitiously bound $Fe^{2+}$. To further define the nature of this component, stearoyl-CoA and ACP were added to an aliquot of the sample used to generate spectrum B of Fig. 4 and the mixture was incubated at 4°C for 1 hr. After this treatment, the intensity of the $Fe^{2+}$ component had decreased from 19% to 7% without generation of mononuclear $Fe^{3+}$. Thus, the addition of substrate promoted catalytic oxidation of the diferrous cluster to the diferric form. Interestingly, the intensity of minority species 2 declined similarly after this treatment, with transfer of absorption into the spectral region of species 1.

**Desaturase Primary Structure.** Although both ribonucleotide reductase and hemerythrin contain diiron-oxo clusters, the x-ray structures clearly show no substantial homology in the primary ligation sphere (14, 15). Similarly, little sequence homology was observed between the desaturase (3) and hemerythrin (15). However, the desaturase subunit contains a pair of the amino acid sequence -(Asp/Glu)-Glu-Xaa-Arg-His- analogous to the known iron-binding sites of ribonucleotide reductase (14) and to the proposed sites of the MMO hydroxylase (13, 16) (see Fig. 6). On the basis of spectroscopic similarities between the desaturase, ribonucleotide reductase, and the MMO hydroxylase reported here, we propose that these sequences represent the iron-binding sites of the desaturase cluster. Since a pair of these sequences are present in each subunit of the desaturase, we further propose

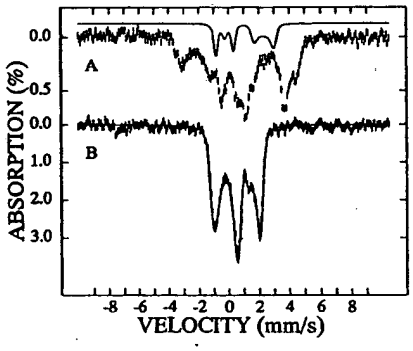



FIG. 5. High-field Mössbauer spectra of the reduced and oxidized desaturase recorded at 4.2 K in a 6.0-T applied field. The samples were the same as those of Fig. 4; the 19% contribution of spectrum A was subtracted from the raw data to provide spectrum B. The solid line above A, representing 20% of the spectral area, was computed assuming that the diferrous cluster had a diamagnetic ground state. The solid line drawn through the data of B was computed using the parameters listed in Table 1 and assuming that species 1 and 2 are diamagnetic. The following values were used for the asymmetry parameter $\eta$ of the quadrupole interactions: $\eta(1) = 1$ and $\eta(2) = 0.2$. These values are uncertain but do not affect the conclusions regarding the diamagnetism of the cluster.

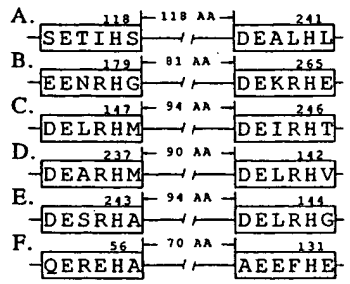



FIG. 6. Primary sequence homologies of diiron-oxo proteins capable of reaction with $O_2$. A, ribonucleotide reductase (14); B, stearoyl-ACP desaturase (3); C, MMO hydroxylase (13); D, phenol hydroxylase (17); E, toluene-4-monooxygenase (18); and F, rubrerythrin (19). Presently, no catalytic reaction with $O_2$ has been reported for rubrerythrin.

that the dimeric desaturase contains a pair of identical diiron-oxo clusters, fully accounting for the analytically and spectroscopically detected iron.

## DISCUSSION

Expression of the recombinant stearoyl-ACP $\Delta^9$ desaturase in *E. coli* has provided sufficient amounts of the desaturase to initiate detailed characterizations of the desaturase active site (20) and the reaction mechanism (21) of oxidative fatty acid desaturation. Here we provide evidence for the presence of a catalytically essential diiron-oxo cluster in the stearoyl $\Delta^9$ desaturase.

The oxidized desaturase exhibits absorbance features between 300 and 350 nm. For methemerythrin, these features have been assigned as ligand-to-metal charge-transfer transitions arising from the Fe—O—Fe bond of a $\mu$-oxo-bridged diiron cluster (9). A similar assignment appears to be appropriate for the desaturase. Hemerythrin also exhibits exogenous ($Cl^-$, $HO^-$, $N_3^-$) ligand-to-metal charge-transfer bands in the region from 400 to 500 nm. It is therefore reasonable to assign the weak absorbance band observed at 475 nm in the oxidized desaturase and the stronger band observed at 450 nm in the presence of azide (see Fig. 2) to such a transition. As $O_2$ binding is likely required during the catalytic cycle, the binding of azide to the desaturase cluster indicates the presence of accessible coordination sites.

The Mössbauer data show that all iron of the $\alpha_2$ desaturase belongs to two clusters of the type found in diiron-oxo proteins. The oxidized clusters have $\Delta E_Q$ and $\delta$ values characteristic of high-spin $Fe^{3+}$ in an environment of oxygen and nitrogen ligands. The absence of low-energy charge-transfer interactions characteristic of thiolate (400–600 nm) coordination further supports this assignment. Mössbauer studies in applied magnetic fields show that the two ferric sites are antiferromagnetically coupled. The strength of the coupling, $J > 60$ $cm^{-1}$, is consistent with, but does not prove, the presence of an oxo bridge between the two ferric ions. The isomer shifts for the diferrous cluster are at the high end of the range of values for diiron-oxo clusters, suggesting that the coordination environment is rich in oxygenous ligands such as carboxylate and water. Since the desaturase is a homodimer and also contains two diiron clusters, it is reasonable to assume that both clusters are equivalent, with each cluster having two inequivalent iron sites in the diferrous form (see Table 1). The presence of paramagnetic hyperfine interactions in a 6.0-T applied field rules out *strong* antiferromagnetic coupling between the ferrous ions but is consistent with either ferromagnetic or *weak* antiferromagnetic coupling.

Upon computer search (GenBank release 74) of primary sequences, only 40 additional occurrences of the iron-binding sequences of Fig. 6 were identified. A majority of these were found in proteins known to contain nonheme iron and to catalyze $O_2$-dependent reactions. Half of these occurrences are in the iron- or manganese-containing superoxide dismutase (22), for which the x-ray structure shows that Asp-156 and His-160 are iron ligands. Isopenicillin N synthase also contains a single copy of this sequence (23). Interestingly, a high-valent iron-oxo intermediate has been proposed for this enzyme (24). For the isopenicillin N synthase from *Cephalosporium acremonium*, Asp-131, Glu-132, and His-135 are contained in the iron-binding sequence, in possible correspondence with the histidine and aspartate ligation proposed from NMR studies (25).

Recently, a diiron-oxo cluster has been observed in rubrerythrin. On the basis of a similar primary sequence analysis, two iron-binding sites were previously identified in rubrerythrin (19). The Mössbauer parameters of the desaturase and rubrerythrin are nearly identical for both redox states and are consistent with the histidine and carboxylate ligation proposed in Fig. 6. We also note that two copies of -(Asp/Glu)-Glu-Xaa-Arg-His- are observed in phenol hydroxylase (17) and toluene-4-monooxygenase (18), suggesting a diiron-oxo cluster may be present in these two oxygenases as well. Finally, for all known sequences of the soluble plant desaturase (six to date), the MMO hydroxylase (13), phenol hydroxylase, and toluene monooxygenase, histidine is always preceded by arginine within the proposed iron-binding sequence, suggesting a possible contribution to oxygenase catalysis as proposed for cytochrome *c* peroxidase (26).

It has become increasingly clear that diiron-oxo clusters are common and catalytically diverse structures. Owing to the requirement to cleave unactivated C—H bonds, oxidative desaturation may involve the generation of a reactive high-valent iron-oxo intermediate of the type proposed for the MMO hydroxylase (10). Structural variations in the ligation sphere may allow partition of the reactivity of the diiron-oxo cluster between oxidative desaturation, hydroxylation, or tyrosine radical formation. The key structural features which modulate the oxidative reactivity of protein-contained diiron-oxo clusters remain to be identified.

We thank A. A. Rangarajan and I. Widders for use of the atomic absorption spectrometer with which the metal determinations were made. J.S. and B.G.F. contributed equally to this work and should therefore both be considered first and corresponding authors. This work was supported by grants from the Department of Agriculture/National Science Foundation/Department of Energy Plant Science Center Program, the National Science Foundation (Grant DCB-891631), and the Department of Energy (Grant DE-AC02-76ER0-13338) to C.S. and by a Grant from the National Institutes of Health (Grant GM-22701) to E.M., and J.S. was a Monsanto Visiting Research Associate at Michigan State University.

1. Ohlrogge, J. B., Kuhn, D. N. & Stumpf, P. K. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1194–1198.
2. Nagai, J. & Bloch, K. (1968) *J. Biol. Chem.* **243**, 4626–4633.
3. Shanklin, J. & Somerville, C. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 2510–2514.
4. Thompson, G. A., Scherer, D. E., Foxall-Van Aken, S., Kenny, J. W., Young, H. L., Shintani, D. K., Kridl, J. C. & Knauf, V. C. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 2578–2582.
5. McKeon, T. A. & Stumpf, P. K. (1982) *J. Biol. Chem.* **257**, 12141–12147.
6. Thiede, M. A., Ozols, J. & Strittmatter, P. (1986) *J. Biol. Chem.* **261**, 13230–13235.
7. Strittmatter, P., Spatz, L., Corcoran, D., Rogers, M. J., Setlow, B. & Redline, R. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 4565–4569.
8. Münck, E. (1978) *Methods Enzymol.* **54**, 346–379.
9. Que, L., Jr. & True, A. E. (1990) in *Progress in Inorganic Chemistry: Bioinorganic Chemistry*, ed., Lippard, S. J. (Wiley, New York), Vol. 38, pp. 97–200.
10. Fox, B. G., Borneman, J. G., Wackett, L. P. & Lipscomb, J. D. (1990) *Biochemistry* **29**, 6419–6427.
11. Elgren, T. E., Lynch, J. B., Juarez-Garcia, C., Münck, E., Sjöberg, B.-M. & Que, L., Jr. (1991) *J. Biol. Chem.* **266**, 19265–19268.
12. Buist, P. H. & Marecak, D. M. (1992) *J. Am. Chem. Soc.* **114**, 5073–5080.
13. Cardy, D. L. N., Laidler, V., Salmond, G. P. C. & Murrell, J. C. (1991) *Mol. Microbiol.* **5**, 335–342.
14. Nordlund, P., Sjöberg, B.-M. & Eklund, H. (1990) *Nature (London)* **345**, 593–598.
15. Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. (1984) *J. Am. Chem. Soc.* **106**, 618–622.
16. Hendrich, M. P., Fox, B. G., Andersson, K. K., Debrunner, P. G. & Lipscomb, J. D. (1992) *J. Biol. Chem.* **267**, 261–269.
17. Nordlund, I., Powlowski, J. & Shingler, V. (1990) *J. Bacteriol.* **172**, 6826–6833.
18. Yen, K.-M., Karl, M. R., Blatt, L. M., Simon, M. J., Winter, R. B., Fausset, P. R., Lu, H. S., Harcourt, A. A. & Chen, K. K. (1991) *J. Bacteriol.* **173**, 5315–5327.
19. Kurtz, D. M., Jr. & Prickril, B. C. (1991) *Biochem. Biophys. Res. Commun.* **181**, 337–341.
20. Schneider, G., Lindqvist, Y., Shanklin, J. & Somerville, C. (1992) *J. Mol. Biol.* **225**, 561–564.
21. van de Loo, F. J., Fox, B. G. & Somerville, C. (1993) in *Lipid Metabolism in Plants*, ed. Moore, T. (CRC, Boca Raton, FL), pp. 91–126.
22. Carlioz, A., Ludwig, M. L., Stallings, W. C., Fee, J. A., Steinman, H. M. & Touati, D. (1988) *J. Biol. Chem.* **263**, 1555–1562.
23. Samson, S. M., Belagaje, R., Blankenship, D. T., Chapman, J. L., Perry, D., Skatrud, P. L., Van Frank, R. M., Abraham, E. P., Baldwin, J. E., Queener, S. W. & Ingolia, T. D. (1985) *Nature (London)* **318**, 191–194.
24. Baldwin, J. E. & Abraham, E. P. (1988) *Nat. Prod. Rep.* **5**, 129–145.
25. Ming, L.-J., Que, L., Jr., Kriauciunas, A., Frolik, C. A. & Chen, V. J. (1990) *Biochemistry* **30**, 11653–11659.
26. Poulos, T. L. & Kraut, J. (1980) *J. Biol. Chem.* **255**, 8199–8205.

# EXHIBIT C

1: Plant Mol Biol. 1993 May;22(2):293-300.



# Isolation of a delta 6-desaturase gene from the cyanobacterium Synechocystis sp. strain PCC 6803 by gain-of-function expressio Anabaena sp. strain PCC 7120.

**Reddy AS, Nuccio ML, Gross LM, Thomas TL.**

Department of Biology, Texas A&M University, College Station 77843.

The enzyme delta 6-desaturase is responsible for the conversion of linoleic a (18:2) to gamma-linolenic acid (18:3 gamma). A cyanobacterial gene encodi delta 6-desaturase was cloned by expression of a Synechocystis genomic cos library in Anabaena, a cyanobacterium lacking delta 6-desaturase. Expressio the Synechocystis delta 6-desaturase gene in Anabaena resulted in the accumulation of gamma-linolenic acid (GLA) and octadecatetraenoic acid (1 The predicted 359 amino acid sequence of the Synechocystis delta 6-desatura shares limited, but significant, sequence similarity with two other reported desaturases. Analysis of three overlapping cosmids revealed a delta 12-desat gene linked to the delta 6-desaturase gene. Expression of Synechocystis delta and delta 12-desaturases in Synechococcus, a cyanobacterium deficient in bo desaturases, resulted in the production of linoleic acid and gamma-linolenic a

PMID: 8389613 [PubMed - indexed for MEDLINE]

# EXHIBIT D

1: Science. 1992 Nov 20;258(5086):1353-5. Related Article

# Map-based cloning of a gene controlling omega-3 fatty acid desaturation in Arabidopsis.

**Arondel V, Lemieux B, Hwang I, Gibson S, Goodman HM, Somerville C**

MSU-DOE Plant Research Laboratory, Michigan State University, East Lans 48824-1312.

A gene from the flowering plant Arabidopsis thaliana that encodes an omega desaturase was cloned on the basis of the genetic map position of a mutation affecting membrane and storage lipid fatty acid composition. Yeast artificial chromosomes covering the genetic locus were identified and used to probe a complementary DNA library. A complementary DNA clone for the desaturas identified and introduced into roots of both wild-type and mutant plants by T plasmid-mediated transformation. Transgenic tissues of both mutant and wil plants had significantly increased amounts of the fatty acid produced by this desaturase.

PMID: 1455229 [PubMed - indexed for MEDLINE]

# EXHIBIT E

**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 1/21, 15/29, 15/82
C07H 15/12

A1

(11) International Publication Number: WO 91/13972

(43) International Publication Date: 19 September 1991 (19.09.91)

(21) International Application Number: PCT/US91/01746

(22) International Filing Date: 14 March 1991 (14.03.91)

(30) Priority data:

| | | |
|---|---|---|
| 494,106 | 16 March 1990 (16.03.90) | US |
| 567,373 | 13 August 1990 (13.08.90) | US |
| 615,784 | 14 November 1990 (14.11.90) | US |

(60) Parent Application or Grant
(63) Related by Continuation
US 494,106 (CIP)
Filed on 16 March 1990 (16.03.90)

(71) Applicant *(for all designated States except US)*: CALGENE, INC. [US/US]; 1920 Fifth Street, Davis, CA 95616 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : THOMPSON, Gregory, A. [US/US]; 5127 Cowell Blvd., Davis, CA 95616 (US). KNAUF, Vic, C. [US/US]; 1013 Hillview Lane, Winters, CA 95694 (US).

(74) Agents: LASSEN, Elizabeth et al.; Calgene, Inc., 1920 Fifth Street, Davis, CA 95616 (US).

(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent), US.

**Published**
*With international search report.*
*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(54) Title: PLANT DESATURASES - COMPOSITIONS AND USES

(57) Abstract

By this invention, compositions and methods of use of plant desaturase enzymes, especially Δ-9 desaturases, are provided. Of special interest are methods and compositions of amino acids and nucleic acid sequences related to biologically active plant desaturases as well as sequences, especially nucleic acid sequences, which are to be used as probes, vectors for transformation or cloning intermediates. Biologically active sequences may be found in a sense or anti-sense orientation as to transcriptional regulatory regions found in various constructs.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| AT | Austria | ES | Spain | MG | Madagascar |
| AU | Australia | FI | Finland | ML | Mali |
| BB | Barbados | FR | France | MN | Mongolia |
| BE | Belgium | GA | Gabon | MR | Mauritania |
| BF | Burkina Faso | GB | United Kingdom | MW | Malawi |
| BG | Bulgaria | GN | Guinea | NL | Netherlands |
| BJ | Benin | GR | Greece | NO | Norway |
| BR | Brazil | HU | Hungary | PL | Poland |
| CA | Canada | IT | Italy | RO | Romania |
| CF | Central African Republic | JP | Japan | SD | Sudan |
| CG | Congo | KP | Democratic People's Republic of Korea | SE | Sweden |
| CH | Switzerland | | | SN | Senegal |
| CI | Côte d'Ivoire | KR | Republic of Korea | SU | Soviet Union |
| CM | Cameroon | LI | Liechtenstein | TD | Chad |
| CS | Czechoslovakia | LK | Sri Lanka | TG | Togo |
| DE | Germany | LU | Luxembourg | US | United States of America |
| DK | Denmark | MC | Monaco | | |

# PLANT DESATURASES - COMPOSITIONS AND USES

This application is a continuation-in-part of USSN 07/494,106 filed on March 16, 1990 and a continuation-in-part of USSN 07/567,373 filed on August 13, 1990 and a continuation-in-part of USSN 07/615,784 filed on November 14, 1990.

## Technical Field

The present invention is directed to desaturase enzymes relevant to fatty acid synthesis in plants, enzymes, amino acid and nucleic acid sequences and methods related thereto, and novel plant entities and/or oils and methods related thereto.

## INTRODUCTION

### Background

Novel vegetable oils compositions and/or improved means to obtain or manipulate fatty acid compositions, from biosynthetic or natural plant sources, are needed. Depending upon the intended oil use, various different oil compositions are desired. For example, edible oil sources containing the minimum possible amounts of saturated fatty acids are desired for dietary reasons and alternatives to current sources of highly saturated oil products, such as tropical oils, are also needed.

One means postulated to obtain such oils and/or modified fatty acid compositions is through the genetic engineering of plants. However, in order to genetically engineer plants one must have in place the means to transfer genetic material to the plant in a stable and heritable manner. Additionally, one must have nucleic acid sequences capable of producing the desired phenotypic result, regulatory regions capable of directing the correct application of such sequences, and the like. Moreover, it

should be appreciated that to produce a desired modified oils phenotype requires that the Fatty Acid Synthetase (FAS) pathway of the plant is modified to the extent that the ratios of reactants are modulated or changed.

Higher plants appear to synthesize fatty acids via a common metabolic pathway in plant plastid organelles (i.e., chloroplasts, proplastids, or other related organelles) as part of the FAS complex. Outside of plastid organelles, fatty acids are incorporated into triglycerides and used in plant membranes and in neutral lipids. In developing seeds, where oils are produced and stored as sources of energy for future use, FAS occurs in proplastids.

The production of fatty acids begins in the plastid with the reaction between Acyl Carrier Protein (ACP) and acetyl-CoA to produce acetyl-ACP. Through a sequence of cylical reactions, the acetyl-ACP is elongated to 16- and 18- carbon fatty acids. The longest chain fatty acids produced by the FAS are 18 carbons long. Monunsaturated fatty acids are also produced in the plastid through the action of a desaturase enzyme.

Common plant fatty acids, such as oleic, linoleic and α-linolenic acids, are the result of sequential desaturation of stearate. The first desaturation step is the desaturation of stearoyl-ACP (C18:0) to form oleoyl-ACP (C18:1) in a reaction often catalyzed by a Δ-9 desaturase, also often referred to as a "stearoyl-ACP desaturase" because of its high activity toward stearate the 18 carbon acyl-ACP. The desaturase enzyme functions to add a double bond at the ninth carbon in accordance with the following reaction (I):

$$\text{Stearoyl-ACP} + \text{ferredoxin(II)} + O_2 + 2H^+ \rightarrow \text{oleoyl-ACP} + \text{ferredoxin (III)} + 2H_2O.$$

Δ-9 desaturases have been studied in partially purified preparations from numerous plant species. Reports indicate that the protein is a dimer, perhaps a homodimer, displaying a molecular weight of 68 kD (±8 kD) by gel-filtration and a molecular weight of 36 kD by SDS-polyacrylamide gel electrophoresis.

In subsequent sequential steps for triglyceride production, polyunsaturated fatty acids may be produced. These desaturations occur outside of the plastid as a result of the action of membrane-bound enzymes. Additional double bonds are added at the twelve position carbon and thereafter, if added, at the 15 position carbon through the action of Δ-12 desaturase and Δ-15 desaturase, respectively.

Obtaining nucleic acid sequences capable of producing a phenotypic result in FAS, desaturation and/or incorporation of fatty acids into a glycerol backbone to produce an oil is subject to various obstacles including but not limited to the identification of metabolic factors of interest, choice and characterization of a protein source with useful kinetic properties, purification of the protein of interest to a level which will allow for its amino acid sequencing, utilizing amino acid sequence data to obtain a nucleic acid sequence capable of use as a probe to retrieve the desired DNA sequence, and the preparation of constructs, transformation and analysis of the resulting plants.

Thus, the identification of enzyme targets and useful plant sources for nucleic acid sequences of such enzyme targets capable of modifying fatty acid compositions are needed. Ideally, an enzyme target will be amenable to one or more applications alone or in combination with other nucleic acid sequences relating to increased/decreased oil production, the ratio of saturated to unsaturated fatty acids in the fatty acid pool, and/or to novel oils compositions as a result of the modifications to the fatty acid pool. Once enzyme target(s) are identified and qualified, quantities of protein and purification protocols are needed for sequencing. Ultimately, useful nucleic acid constructs having the necessary elements to provide a phenotypic modification and plants containing such constructs are needed.

Relevant Literature

A 200-fold purification of *Carthamus tinctorius* ("safflower") stearoyl-ACP desaturase was reported by McKeon & Stumpf in 1982, following the first publication of their protocol in 1981. McKeon, T. & Stumpf, P. *J.Biol.Chem.* (1982) *257*:12141-12147; McKeon, T. & Stumpf, P. *Methods in Enzymol.* (1981) *71*:275-281.

## BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 provides amino acid sequence of fragments relating to *C. tinctorius* desaturase. Fragments F1 through F11 are also provided in the sequence listing as SEQ ID NO: 1 through SEQ ID NO: 11, respectively. Each fragment represents a synthesis of sequence information from peptides originating from different digests which have been matched and aligned. In positions where there are two amino acids indicated, the top one corresponds to that found in the translation of the cDNA; the lower one was detected either as a second signal at the same position of one of the sequenced peptides, or as a single unambiguous signal found in one or more of the overlapping peptides comprising the fragment. Residues in F9 shown in lower case letters represent positions where the called sequence does not agree with that predicted from the cDNA, but where the amino acid assignment is tentative because of the presence of a contaminating peptide. The standard one letter code for amino acid residues has been used. X represents a position where no signal was detectable, and which could be a modified residue. F1 corresponds to the N-terminal sequence of the mature protein. The underlined region in F2 is the sequence used in designing PCR primers for probe synthesis.

Fig. 2 provides a cDNA sequence (SEQ ID NO: 12) and the corresponding translational peptide sequence (SEQ ID NO: 13) derived from *C. tinctorius* desaturase. The cDNA sequence includes both the plastid transit peptide encoding sequence and the sequence encoding the mature protein.

Fig. 3 provides cDNA sequence of *Ricinus communis* desaturase. Fig. 3A provides preliminary partial cDNA sequence of a 1.7 kb clone of *R. communis* desaturase (SEQ ID NO: 14). The sequence is from the 5' end of the clone. Fig. 3B provides the complete cDNA sequence of the approximately 1.7 kb clone (SEQ ID NO: 15) and the corresponding translational peptide sequence (SEQ ID NO: 16).

Fig. 4 provides sequence of *Brassica campestris* desaturase. Fig. 4A represents partial DNA sequence of a 1.6 kb clone pCGN3235 (SEQ ID NO: 17), from the 5' end of the clone. Fig. 4B represents partial DNA sequence of a 1.2 kb clone, pCGN3236, from the 5' end of the clone (SEQ ID NO: 18). Initial sequence for the 3' ends of the two *B. campestris* desaturase clones indicates that pCGN3236 is a shorter cDNA for the same clone as pCGN3235. Fig. 4C provides complete cDNA sequence of *B. campestris* desaturase above, pCGN3235 (SEQ ID NO: 19) and the corresponding translational peptide sequence (SEQ ID NO: 20).

Fig. 5 provides preliminary partial cDNA sequence of *Simmondsia chinensis* desaturase (SEQ ID NO: 43). The translated amino acid sequence is also shown.

Fig. 6 shows the design of forward and reverse primers (SEQ ID NO: 21 through SEQ ID NO: 26) used in polymerase chain reaction (PCR) from the sequence of *C. tinctorius* desaturase peptide "Fragment F2" (SEQ ID NO: 2).

Fig. 7 provides maps of desaturase cDNA clones showing selected restriction enzyme sites. Fig. 7A represents a *C. tinctorius* clone, Fig. 7B represents a *R. communis* clone, and Fig. 7C represents a *B. campestris* clone.

Fig. 8 provides approximately 3.4 kb of genomic sequence of Bce4 (SEQ ID NO: 27).

Fig. 9 provides approximately 4 kb of genomic sequence of Bcg 4-4 ACP sequence (SEQ ID NO: 28).

Fig. 10 provides a restriction map of cloned λCGN 1-2 showing the entire napin coding region sequence as well as extensive 5' upstream and 3' downstream sequences (SEQ ID NO: 29).

## SUMMARY OF THE INVENTION

By this invention, compositions and methods of use of plant desaturase enzymes, especially Δ-9 desaturases, are provided. Of special interest are methods and compositions of amino acids and nucleic acid sequences related to biologically active plant desaturases as well as sequences, especially nucleic acid sequences, which are to be used as probes, vectors for transformation or cloning intermediates. Biologically active sequences may be found in a sense or anti-sense orientation as to transcriptional regulatory regions found in various constructs.

A first aspect of this invention relates to C. tinctorius Δ-9 desaturase substantially free of seed storage protein. Amino acid sequence of this desaturase is provided in Fig. 2 and as SEQ ID NO: 13.

DNA sequence of *C. tinctorius* desaturase gene (SEQ ID NO: 12) is provided, as well as DNA sequences of desaturase genes from a *Ricinus* (SEQ ID NO: 14 and SEQ ID NO: 15) a *Brassica* (SEQ ID NO: 17 through SEQ ID NO: 19) and a *Simmondsia* (SEQ ID NO: 43) plant.

In yet a different embodiment of this invention, plant desaturase cDNA of at least 10 nucleotides or preferably at least 20 nucleotides and more preferably still at least 50 nucleotides, known or homologously related to known Δ-9 desaturase(s) is also provided. The cDNA encoding precursor desaturase or, alternatively, biologically active, mature desaturase is provided herein.

Methods to use nucleic acid sequences to obtain other plant desaturases are also provided. Thus, a plant desaturase may be obtained by the steps of contacting a nucleic acid sequence probe comprising nucleotides of a known desaturase sequence and recovery of DNA sequences encoding plant desaturase having hybridized with the probe.

This invention also relates to methods for obtaining plant Δ-9 desaturase by contacting an antibody specific to a known desaturase, such as *C. tinctorius* stearoyl-ACP

desaturase, with a candidate plant stearoyl-ACP desaturase under conditions conducive to the formation of an antigen:antibody immunocomplex and the recovery of the candidate plant stearoyl-ACP desaturase which reacts thereto.

In a further aspect of this invention DNA constructs comprising a first DNA sequence encoding a plant desaturase and a second DNA sequence which is not naturally found joined to said plant desaturase are provided. This invention also relates to the presence of such constructs in host cells, especially plant host cells. In yet a different aspect, this invention relates to transgenic host cells which have an expressed desaturase therein.

Constructs of this invention may contain, in the 5' to 3' direction of transcription, a transcription initiation control regulatory region capable of promoting transcription in a host cell and a DNA sequence encoding plant desaturase. Transcription initiation control regulatory regions capable of expression in prokatyotic or eukaryotic host cells are provided. Most preferred are transcription initiation control regions capable of expression in plant cells, and more preferred are transcription and translation initiation regions preferentially expressed in plant cells during the period of lipid accumulation. The DNA sequence encoding plant desaturase of this invention may be found in either the sense or anti-sense orientation to the transcription initiation control region.

Specific constructs, expression cassettes having in the 5' to 3' direction of transcription, a transcription and translation initiation control regulatory region comprising sequence immediately 5' to a structural gene preferentially expressed in plant seed during lipid accumulation, a DNA sequence encoding desaturase, and sequence 3' to the structural gene are also provided. The construct may preferably contain DNA sequences encoding plant desaturase obtainable (included obtained) from *Carthamus*, *Rininus*, *Brassica* or *Simmondsia* Δ-9 desaturase

genes. Transcription and translation initiation control regulatory regions are preferentially obtained from structural genes preferentially expressed in plant embryo tissue such as napin, seed-ACP or Bce-4.

By this invention, methods and constructs to inhibit the production of endogenous desaturase are also provided. For example, an anti-sense construct comprising, in the 5' to 3' direction of transcription, a transcription initiation control regulatory region functional in a plant cell, and an anti-sense DNA sequence encoding a portion of a plant Δ-9 desaturase may be integrated into a plant host cell to decrease desaturase levels.

In yet a different embodiment, this invention is directed to a method of producing plant desaturase in a host cell comprising the steps of growing a host cell comprising an expression cassette, which would contain in the direction of transcription, a) a transcription and translation initiation region functional in said host cell, b) the DNA sequence encoding a plant desaturase in reading frame with said initiation region, and c) and a transcript termination region functional in said host cell, under conditions which will promote the expression of the plant desaturase. Cells containing a plant desaturase as a result of the production of the plant desaturase encoding sequence and also contemplated herein.

By this invention, a method of modifying fatty acid composition in a host plant cell from a given level of fatty acid saturation to a different level of fatty acid saturation is provided by growing a host plant cell having integrated into its genome a recombinant DNA sequence encoding a plant desaturase in either a sense or anti-sense orientation under control of regulatory elements functional in said plant cell during lipid accumulation under conditions which will promote the activity of said regulatory elements. Plant cells having such a modified level of fatty acid saturation are also contemplated hereunder. Oilseeds having such a modified level of fatty

acid saturation and oils produced from such oilseeds are further provided.

## DETAILED DESCRIPTION OF THE INVENTION

A plant desaturase of this invention includes any sequence of amino acids, such as a protein, polypeptide, or peptide fragment, obtainable from a plant source which is capable of catalyzing the insertion of a first double bond into a fatty acyl-ACP moiety in a plant host cell, i.e., *in vivo*, or in a plant cell-like environment, i.e. *in vitro*. "A plant cell-like environment" means that any necessary conditions are available in an environment (i.e., such factors as temperatures, pH, lack of inhibiting substances) which will permit the enzyme to function In particular, this invention relates to enzymes which add such a first double bond at the ninth carbon position in a fatty acyl-ACP chain. There may be similar plant desaturase enzymes of this invention with different specificities, such as the Δ-12 desaturase of carrot.

Nucleotide sequences encoding desaturases may be obtained from natural sources or be partially or wholly artificially synthesized. They may directly correspond to a desaturase endogenous to a natural plant source or contain modified amino acid sequences, such as sequences which have been mutated, truncated, increased or the like. Desaturases may be obtained by a variety of methods, including but not limited to, partial or homogenous purification of plant extracts, protein modeling, nucleic acid probes, antibody preparations and sequence comparisons. Typically a plant desaturase will be derived in whole or in part from a natural plant source.

Of special interest are Δ-9 desaturases which are obtainable, including those with are obtained, from *Cartharmus*, *Ricinus*, *Simmondsia*, or *Brassica*, for example *C. tinctorius*, *R. communis*, *S. chinensis* and *B. campestris*, respectively, or from plant desaturases which are obtainable through the use of these sequences. "Obtainable" refers to those desaturases which have

sufficiently similar sequences to that of the native sequences provided herein to provide a biologically active desaturase.

Once a DNA sequence which encodes a desaturase is obtained, it may be employed as a gene of interest in a nucleic acid construct or in probes in accordance with this invention. A desaturase may be produced in host cells for harvest or as a means of effecting a contact between the desaturase and its substrate. Constructs may be designed to produce desaturase in either prokaryotic or eukaryotic cells. Plant cells containing recombinant constructs encoding biologically active desaturase sequences, both expression and anti-sense constructs, as well as plants and cells containing modified levels of desaturase proteins are of special interest. For use in a plant cell, constructs may be designed which will effect an increase or a decrease in amount of endogenous desaturase available to a plant cell transformed with such a construct.

Where the target gene encodes an enzyme, such as a plant desaturase, which is already present in the host plant, there are inherent difficulties in analyzing mRNA, engineered protein or enzyme activity, and modified fatty acid composition or oil content in plant cells, especially in developing seeds; each of which can be evidence of biological activity. This is because the levels of the message, enzyme and various fatty acid species are changing rapidly during the stage where measurements are often made, and thus it can be difficult to discriminate between changes brought about by the presence of the foreign gene and those brought about by natural developmental changes in the seed. Where an expressed Δ-9 desaturase DNA sequence is derived from a plant species heterologous to the plant host into which the sequence is introduced and has a distinguishable DNA sequence, it is often possible to specifically probe for expression of the foreign gene with oligonucleotides complimentary to unique sequences of the inserted DNA/RNA. And, if the foreign gene codes for a protein with slightly different protein sequence, it may be

possible to obtain antibodies which recognize unique epitopes on the engineered protein. Such antibodies can be obtained by mixing the antiserum to the foreign protein with extract from the host plant, or with extracts containing the host plant enzyme. For example, one can isolate antibodies uniquely specific to a *C. tinctorius* Δ-9 desaturase by mixing antiserum to the desaturase with an extract containing a *Brassica* Δ-9 desaturase. Such an approach will allow the detection of *C. tinctorius* desaturase in *Brassica* plants transformed with the *C. tinctorius* desaturase gene. In plants expressing an endogenous gene in an antisense orientation, the problem is slightly different. In this case, there are no specific reagents to measure expression of a foreign protein. However, one is attempting to measure a decrease in an enzyme activity that normally is increasing during development. This makes detection of expression a simpler matter. In the final seed maturation phase, enzyme activities encoded by genes affecting oil composition usually disappear and cannot be detected in final mature seed. Analysis of the fatty acid content may be preformed by any manner known to those skilled in the art, including gas chromatography, for example.

By increasing the amount of desaturase available in the plant cell, an increased percentage of unsaturated fatty acids may be provided; by decreasing the amount of desaturase, an increased percentage of saturated fatty acids may be provided. (Modifications in the pool of fatty acids available for incorporation into triglycerides may likewise affect the composition of oils in the plant cell.) Thus, an increased expression of desaturase in a plant cell may result in increased proportion of fatty acids, such as one or more of palmitoleate (C16:1), oleate (C18:1), linoleate (C18:2) and linolenate (C18:3) are expected. Of special interest is the production of triglycerides having increased levels of oleate. Using anti-sense technology, alternatively, a decrease in the amount of desaturase available to the plant cell is expected, resulting in a

higher percentage of saturates such as one or more of laurate (C12:0), myristate (C14:0), palmitate (C16:0), stearate (C18:0), arachidate (C20:0), behemate (C22:0) and lignocerate (C24:0). Of special interest is the production of triglycerides having increased levels of stearate or palmitate and stearate. In addition, the production of a variety of ranges of such saturates is desired. Thus, plant cells having lower and higher levels of stearate fatty acids are contemplated. For example, fatty acid compositions, including oils, having a 10% level of stearate as well as compositions designed to have up to a 60% level of stearate or other such modified fatty acid(s) composition are contemplated.

The modification of fatty acid compositions may also affect the fluidity of plant membranes. Different lipid concentrations have been observed in cold-hardened plants, for example. By this invention, one may be capable of introducing traits which will lend to chill tolerance. Constitutive or temperature inducible transcription initiation regulatory control regions may have special applications for such uses.

Other applications for use of cells or plants producing desaturase may also be found. For example, potential herbicidal agents selective for plant desaturase may be obtained through screening to ultimately provide environmentally safe herbicide products. The plant desaturase can also be used in conjunction with chloroplast lysates to enhance the production and/or modify the composition of the fatty acids prepared *in vitro*. The desaturase can also be used for studying the mechanism of fatty acid formation in plants and bacteria. For these applications, constitutive promoters may find the best use.

Constructs which contain elements to provide the transcription and translation of a nucleic acid sequence of interest in a host cell are "expression cassettes". Depending upon the host, the regulatory regions will vary, including regions from structural genes from viruses, plasmid or chromosomal genes, or the like. For expression

in prokaryotic or eukaryotic microorganisms, particularly unicellular hosts, a wide variety of constitutive or regulatable promoters may be employed. Among transcriptional initiation regions which have been described are regions from bacterial and yeast hosts, such as *E. coli, B. subtilis, Saccharomyces cerevisiae,* including genes such as β-galactosidase, T7 polymerase, trp E and the like.

A recombinant construct for expression of desaturase in a plant cell ("expression cassette") will include, in the 5' to 3' direction of transcription, a transcription and translation initiation control regulatory region (the transcriptional and translational initiation regions together often also known as a "promoter") functional in a plant cell, a nucleic acid sequence encoding a plant desaturase, and a transcription termination region. Numerous transcription initiation regions are available which provide for a wide variety of constitutive or regulatable, e.g., inducible, transcription of the desaturase structural gene. Among transcriptional initiation regions used for plants are such regions associated with cauliflower mosaic viruses (35S, 19S), and structural genes such as for nopaline synthase or mannopine synthase or napin and ACP promoters, etc. The transcription/translation initiation regions corresponding to such structural genes are found immediately 5' upstream to the respective start codons. Thus, depending upon the intended use, different promoters may be desired.

Of special interest in this invention are the use of promoters which are capable of preferentially expressing the desaturase in seed tissue, in particular, at early stages of seed oil formation. Examples of such seed-specific promoters include the region immediately 5' upstream of napin or seed ACP genes, such as described in co-pending USSN 147,781, and the Bce-4 gene such as described in co-pending USSN 494,722. Alternatively, the use of the 5' regulatory region associated with an endogenous plant desaturase structural gene and/or the

transcription termination regions found immediately 3' downstream to the gene, may often be desired.

In addition, for some applications, use of more than one promoter may be desired. For example, one may design a dual promoter expression cassette each promoter having a desaturase sequence under its regulatory control. For example, the combination of an ACP and napin cassette could be useful for increased production of desaturase in a seed-specific fashion over a longer period of time than either individually.

To decrease the amount of desaturase found in a plant host cell, anti-sense constructs may be prepared and then inserted into the plant cell. By "anti-sense" is meant a DNA sequence in the 5' to 3' direction of transcription in relation to the transcription initiation region, which encodes a sequence complementary to the sequence of a native desaturase. It is preferred that an anti-sense plant desaturase sequence be complementary to a plant desaturase gene indigenous to the plant host. Sequences found in an anti-sense orientation may be found in constructs providing for transcription or transcription and translation of the DNA sequence encoding the desaturase, including expression cassettes. Constructs having more than one desaturase sequence under the control of more than one promoter or transcription initiation region may also be employed with desaturase constructs. Various transcription initiation regions may be employed. One of ordinary skill in the art can readily determine suitable regulatory regions. Care may be necessary in selecting transcription initiation regions to avoid decreasing desaturase activity in plant cells other than oilseed tissues. Any transcription initiation region capable of directing expression in a plant host which causes initiation of adequate levels of transcription selectively in storage tissues during seed development for example, should be sufficient. As such, seed specific promoters may be desired. Other manners of decreasing the amount of endogenous plant desaturase, such as ribozymes or the

screening of plant cells transformed with constructs for rare events containing sense sequences which in fact act to decrease desaturase expression, are also contemplated herein. Other analogous methods may be applied by those of ordinary skill in the art.

By careful selection of plants, transformants having particular oils profiles may be obtained. This may in part depend upon the qualities of the transcription initiation region(s) employed or may be a result of culling transformation events to exploit the variabilities of expression observed.

In order to obtain the nucleic acid sequences encoding *C. tinctorius* desaturase, a protein preparation free of a major albumin-type contaminant is required. As demonstrated more fully in the Examples, the protocols of McKeon and Stumpf, *supra*, result in a preparation contaminated with a seed storage protein. Removal of the protein contaminant may be effected by application of a reverse-phase HPLC, or alternatively, by application of a reduction and alkylation step followed by electrophoresis and blotting, for example. Other purification methods may be employed as well, now that the presence of the contaminant is confirmed and various properties thereof described. Once the purified desaturase is obtained it may be used to obtain the corresponding amino acid and/or nucleic acid sequences thereto in accordance with methods familiar to those skilled in the art. Approximately 90% of the total amino acid sequence of the *C. tinctorius* desaturase is provided in Fig. 1 and in SEQ ID NOS: 1-11. The desaturase produced in accordance with the subject invention can be used in preparing antibodies for assays for detecting plant desaturase from other sources.

A nucleic acid sequence of this invention may include genomic or cDNA sequence and mRNA. A cDNA sequence may or may not contain pre-processing sequences, such as transit peptide sequences. Transit peptide sequences facilitate the delivery of the protein to a given organelle and are

cleaved from the amino acid moiety upon entry into the organelle, releasing the "mature" sequence.

In Fig. 2 and SEQ ID NO: 13, the sequence of the *C. tinctorius* desaturase precursor protein is provided; both the transit peptide and mature protein sequence are shown. Also provided in this invention are cDNA sequences relating to *R. communis* desaturase (Fig.3 and SEQ ID NOS: 14-15), *B. campestris* desaturase (Fig. 4 and SEQ ID NOS: 17-19) and *S. chinesis* (Fig. 5 and SEQ ID NOS: 43).

The use of the precursor cDNA sequence is preferred in desaturase expression cassettes. In addition, desaturase transit peptide sequences may be employed to translocate other proteins of interest to plastid organelles for a variety of uses, including the modulation of other enzymes related to the FAS pathway. See, European Patent Application Publication No. 189,707.

As described in more detail below, the complete genomic sequence of a desaturase may be obtained by the screening of a genomic library with a desaturase cDNA probe and isolating those sequences which regulate expression in seed tissue. In this manner, the transcription, translation initiation regions and/or transcript termination regions of the desaturase may be obtained for use in a variety of DNA constructs, with or without the respective desaturase structural gene.

Other nucleic acid sequences "homologous" or "related" to DNA sequences encoding other desaturases are also provided. "Homologous" or "related" includes those nucleic acid sequences which are identical or conservatively substituted as compared to the exemplified *C. tinctorius, R. communis, S. chinesis* or *B. campestris* desaturase sequences of this invention or a plant desaturase which has in turn been obtained from a plant desaturase of this invention. By conservatively substituted is meant that codon substitutions encode the same amino acid, as a result of the degeneracy of the DNA code, or that a different amino acid having similar properties to the original amino acid is substituted. One skilled in the art will readily

recognize that antibody preparations, nucleic acid probes (DNA and RNA) sequences encoding and the like may be prepared and used to screen and recover desaturase from other plant sources. Typically, nucleic acid probes are labeled to allow detection, preferably with radioactivity although enzymes or other methods may also be used. For immunological screening methods, antibody preparations either monoclonal or polyclonal are utilized. Polyclonal antibodies, although less specific, typically are more useful in gene isolation. For detection, the antibody is labeled using radioactivity or any one of a variety of second antibody/enzyme conjugate systems that are commercially available. Examples of some of the available antibody detection systems are described by Oberfilder (*Focus* (1989) BRL Life Technologies, Inc., *11*:1-5).

A "homologous" or "related" nucleic acid sequence will show at least about 60% homology, and more preferably at least about 70% homology, between the known desaturase sequence and the desired candidate plant desaturase of interest, excluding any deletions which may be present. Homology is determined upon comparison of sequence information, nucleic acid or amino acid, or through hybridization reactions. Amino acid sequences are considered homologous by as little as 25% sequence identity between the two complete mature proteins. (*See generally*, Doolittle, R.F., of URFS and ORFS, University Science Books, CA, 1986.)

Oligonucleotide probes can be considerably shorter than the entire sequence, but should be at least about 10, preferably at least about 15, more preferably at least 20 nucleotides in length. When shorter length regions are used for comparison, a higher degree of sequence identity is required than for longer sequences. Shorter probes are often particularly useful for polymerase chain reactions (PCR), especially when highly conserved sequences can be identified. (*See*, Gould, *et al.*, *PNAS USA* (1989) *86*:1934-1938.) Longer oligonucleotides are also useful, up to the full length of the gene encoding the polypeptide of

interest. When longer nucleic acid fragments are employed (>100 bp) as probes, especially when using complete or large cDNA sequences, one would screen with low stringencies (for example 40-50°C below the melting temperature of the probe) in order to obtain signal from the target sample with 20-50% deviation, i.e., homologous sequences. (*See*, Beltz, *et al.*, *Methods in Enzymology* (1983) *100*:266-285.) Both DNA and RNA probes can be used.

A genomic library prepared from the plant source of interest may be probed with conserved sequences from a known desaturase to identify homologously related sequences. Use of the entire cDNA may be employed if shorter probe sequences are not identified. Positive clones are then analyzed by restriction enzyme digestion and/or sequencing. When a genomic library is used, one or more sequences may be identified providing both the coding region, as well as the transcriptional regulatory elements of the desaturase gene from such plant source. In this general manner, one or more sequences may be identified providing both the coding region, as well as the transcriptional regulatory elements of the desaturase gene from such plant source.

In use, probes are typically labeled in a detectable manner (for example with $^{32}P$-labeled or biotinylated nucleotides) and are incubated with single-stranded DNA or RNA from the plant source in which the gene is sought, although unlabeled oligonucleotides are also useful. Hybridization is detected by means of the label after single-stranded and double-stranded (hybridized) DNA or DNA/RNA have been separated, typically using nitrocellulose paper or nylon membranes. Hybridization techniques suitable for use with oligonucleotides are well known to those skilled in the art. Thus, plant desaturase genes may be isolated by various techniques from any convenient plant. Plant desaturase of developing seed obtained from other oilseed plants, such as soybean, coconut, oilseed rape, sunflower, oil palm, peanut, cocoa, cotton, corn and the like are desired as well as from non-traditional oil

sources, including but not limited to spinach chloroplast, avocado mesocarp, cuphea, California Bay, cucumber, carrot, meadowfoam, *Oenothera* and *Euglena gracillis*.

Once the desired plant desaturase sequence is obtained, it may be manipulated in a variety of ways. Where the sequence involves non-coding flanking regions, the flanking regions may be subjected to resection, mutagenesis, etc. Thus, transitions, transversions, deletions, and insertions may be performed on the naturally occurring sequence. In addition, all or part of the sequence may be synthesized, where one or more codons may be modified to provide for a modified amino acid sequence, or one or more codon mutations may be introduced to provide for a convenient restriction site or other purpose involved with construction or expression. The structural gene may be further modified by employing synthetic adapters, linkers to introduce one or more convenient restriction sites, or the like.

Recombinant constructs containing a nucleic acid sequence encoding a desaturase of this invention may be combined with other, i.e. "heterologous," DNA sequences in a variety of ways. By heterologous DNA sequences is meant any DNA sequence which is not naturally found joined to the native desaturase, including combinations of DNA sequences from the same plant of the plant desaturase which are not naturally found joined together. In a preferred embodiment, the DNA sequence encoding a plant desaturase is combined in a DNA construct having, in the 5' to 3' direction of transcription, a transcription initiation control region capable of promoting transcription in a host cell, and a DNA sequence encoding a desaturase in either a sense or anti-sense orientation. As described in more detail elsewhere, a variety of regulatory control regions containing transcriptional or transcriptional and translational regions may be employed, including all or part of the non-coding regions of the plant desaturase.

The open reading frame coding for the plant desaturase or functional fragment thereof will be joined at its 5' end

to a transcription initiation regulatory control region. In some instances, such as modulation of plant desaturase via a desaturase in an anti-sense orientation, a transcription initiation region or transcription/ translation initiation region may be used. In embodiments wherein the expression of the desaturase protein is desired in a plant host, a transcription/ translation initiation regulatory region, is needed. Additionally, modified promoters, i.e., having transcription initiation regions derived from one gene source and translation initiation regions derived from a different gene source or enhanced promoters, such as double 35S CaMV promoters, may be employed for some applications.

As described above, of particular interest are those 5' upstream non-coding regions which are obtained from genes regulated during seed maturation, particularly those preferentially expressed in plant embryo tissue, such as ACP-and napin-derived transcription initiation control regions. Such regulatory regions are active during lipid accumulation and therefore offer potential for greater control and/or effectiveness to modify the production of plant desaturase and/or modification of the fatty acid composition. Especially of interest are transcription initiation regions which are preferentially expressed in seed tissue, i.e., which are undetectable in other plant parts. For this purpose, the transcript initiation region of acyl carrier protein isolated from *B. campestris* seed and designated as "Bcg 4-4" and an unidentified gene isolated from *B. campestris* seed and designated as "Bce-4" are also of substantial interest.

Briefly, Bce4 is found in immature embryo tissue at least as early as 11 days after anthesis (flowering), peaking about 6 to 8 days later or 17-19 days post-anthesis, and becoming undetectable by 35 days post-anthesis. The timing of expression of the Bce4 gene closely follows that of lipid accumulation in seed tissue. Bce4 is primarily detected in seed embryo tissue and to a lesser extent found in the seed coat. Bce4 has not been

detected in other plant tissues tested, root, stem and leaves.

Approximately 3.4 kb genomic sequence of Bce4 is provided in Fig. 8 and as SEQ ID NO: 27, including about 1 kb 5' to the structural gene, about 0.3 kb of the Bce4 coding gene sequence, and about 2.1 kb of the non-coding regulatory 3' sequence. Bce4 transcript initiation regions will contain at least 1 kb and more preferably about 5 to about 7.5 kb of sequence immediately 5' to the Bce4 structural gene.

The Bcg 4-4 ACP message presents a similar expression profile to that of Bce4 and, therefore, also corresponds to lipid accumulation in the seed tissue. Bcg 4-4 is not found in the seed coat and may show some differences in expression level, as compared to Bce4, when the Bcg 4-4 5' non-coding sequence is used to regulate transcription or transcription and translation of a plant stearoyl-ACP desaturase of this invention. Genomic sequence of Bcg 4-4 is provided in Fig. 9 and as SEQ ID NO: 28, including about 1.5 kb 5' to the structural gene, about 1.2 kb of the Bcg 4-4 (ACP) structural gene sequence, and about 1.3 kb of the non-coding regulatory 3' sequence.

The napin 1-2 message is found in early seed development and thus, also offers regulatory regions which can offer preferential transcriptional regulation of a desired DNA sequence of interest such as the plant desaturase DNA sequence of this invention during lipid accumulation. Napins are one of the two classes of storage proteins synthesized in developing *Brassica* embryos (Bhatty, *et al., Can J. Biochem.* (1968) *46*:1191-1197) and have been used to direct tissue-specific expression when reintroduced into the *Brassica* genome (Radke, *et al., Theor. Appl. Genet.* (1988) *75*:685-694). Genomic sequence of napin 1-2 is provided in Fig. 10 and as SEQ ID NO: 29, including about 1.7 kb 5' to the structural gene and about 1.3 kb of the non-coding regulatory 3' sequence

Regulatory transcript termination regions may be provided in DNA constructs of this invention as well.

Transcript termination regions may be provided by the DNA sequence encoding the plant desaturase or a convenient transcription termination region derived from a different gene source, especially the transcript termination region which is naturally associated with the transcript initiation region. The transcript termination region will contain at least about 1 kb, preferably about 3 kb of sequence 3' to the structural gene from which the termination region is derived.

In developing the DNA construct, the various components of the construct or fragments thereof will normally be inserted into a convenient cloning vector which is capable of replication in a bacterial host, e.g., *E. coli*. Numerous vectors exist that have been described in the literature. After each cloning, the plasmid may be isolated and subjected to further manipulation, such as restriction, insertion of new fragments, ligation, deletion, insertion, resection, etc., so as to tailor the components of the desired sequence. Once the construct has been completed, it may then be transferred to an appropriate vector for further manipulation in accordance with the manner of transformation of the host cell.

Normally, included with the DNA construct will be a structural gene having the necessary regulatory regions for expression in a host and providing for selection of transformant cells. The gene may provide for resistance to a cytotoxic agent, e.g. antibiotic, heavy metal, toxin, etc., complementation providing prototrophy to an auxotrophic host, viral immunity or the like. Depending upon the number of different host species into which the expression construct or components thereof are introduced, one or more markers may be employed, where different conditions for selection are used for the different hosts.

The manner in which the DNA construct is introduced into the plant host is not critical to this invention. Any method which provides for efficient transformation may be employed. Various methods for plant cell transformation include the use of Ti- or Ri-plasmids, microinjection,

electroporation, liposome fusion, DNA bombardment or the like. In many instances, it will be desirable to have the construct bordered on one or both sides by T-DNA, particularly having the left and right borders, more particularly the right border. This is particularly useful when the construct uses *A. tumefaciens* or *A. rhizogenes* as a mode for transformation, although the T-DNA borders may find use with other modes of transformation.

Where *Agrobacterium* is used for plant cell transformation, a vector may be used which may be introduced into the *Agrobacterium* host for homologous recombination with T-DNA or the Ti- or Ri-plasmid present in the *Agrobacterium* host. The Ti- or Ri-plasmid containing the T-DNA for recombination may be armed (capable of causing gall formation) or disarmed (incapable of causing gall formation), the latter being permissible, so long as the *vir* genes are present in the transformed *Agrobacterium* host. The armed plasmid can give a mixture of normal plant cell and gall.

A preferred method for the use of *Agrobacterium* as the vehicle for transformation of plant cells employs a vector having a broad host range replication system, at least one T-DNA boundary and the DNA sequence or sequences of interest. Commonly used vectors include pRK2 or derivatives thereof. See, for example, Ditta *et al.*, *PNAS USA*, (1980) 77:7347-7351 and EPA 0 120 515, which are incorporated herein by reference. Normally, the vector will be free of genes coding for opines, oncogenes and *vir*-genes. Included with the expression construct and the T-DNA will be one or more markers, which allow for selection of transformed *Agrobacterium* and transformed plant cells. A number of markers have been developed for use with plant cells, such as resistance to chloramphenicol, the aminoglycoside G418, hygromycin, or the like. The particular marker employed is not essential to this invention, one or another marker being preferred depending on the particular host and the manner of construction.

The expression constructs may be employed with a wide variety of plant life, particularly plant life involved in the production of vegetable oils. These plants include, but are not limited to rapeseed, sunflower, *C. tinctorius*, cotton, *Cuphea*, peanut, soybean, oil palm and corn. Anti-sense constructs may be employed in such plants which share complementarity between the endogenous sequence and the anti-sense desaturase. Of special interest is the use of an anti-sense construct having a *B. campestris* desaturase in rapeseed, including *B. campestris* and *B. napus*.

For transformation of plant cells using *Agrobacterium*, explants may be combined and incubated with the transformed *Agrobacterium* for sufficient time for transformation, the bacteria killed, and the plant cells cultured in an appropriate selective medium. Once callus forms, shoot formation can be encouraged by employing the appropriate plant hormones in accordance with known methods and the shoots transferred to rooting medium for regeneration of plants. The plants may then be grown to seed and the seed used to establish repetitive generations and for isolation of vegetable oils compositions. A variety of stable genetic lines having fixed levels of saturation may be obtained and integrated into a traditional breeding program. Hemizygous and heterozygous lines or homozygous lines may demonstrate different useful properties for oil production and/or breeding. For example, saturation levels may be increased up to 2-fold by the development of homozygous plants as compared with heterozygous (including hemizygous) plants.

The invention now being generally described, it will be more readily understood by reference to the following examples which are included for purposes of illustration only and are not intended to limit the present invention.

## EXAMPLES

### MATERIALS

Commercially available biological chemicals and chromatographic materials, including BSA, catalase (twice

crystalized from bovine liver), spinach ferredoxin, ferredoxin-$NADP^+$ oxidoreductase (spinach leaf), NADPH, unlabeled fatty acids, DEAE-cellulose (Whatman DE-52) CNBr-activated Sepharose 4B, and octyl-Sepharose, and Reactive Blue Agarose are from Sigma (St. Louis, MO). Triethylamine, trichloroacetic acid, guanidine-HCl, and hydrazine-hydrate are also from Sigma. Proteolytic enzymes, including endoproteinases lysC, gluC, and aspN are sequencing grade enzymes obtained from Boehringer Mannheim (Indianapolis, IN). Organic solvents, including acetone, acetonitrile, methanol, ether and petroleum ether are purchased from J.T. Baker (Phillipsburg, NJ); concentrated acids and sodium sulfate are also from J.T. Baker (Phillipsburg, NJ). HPLC grade acetonitrile and trifluoracetic acid (TFA) are obtained from Burdick and Jackson (Muskegon, MI), and from Applied Biosystems (Foster City, CA), respectively. Radiochemicals, including [9,10(n)-$^3$H] oleic acid (10mCi/μmol) and [$^3$H]-iodoacetic acid (64Ci/mol) are from New England Nuclear (Boston, MA). Phenacyl-8 Reagent (bromoacetophenone with a crown ether catalyst) used to prepare phenacyl esters of the fatty acids for analysis are from Pierce (Rockford, IL). C18 reversed-phase thin-layer chromatography plates are from Whatman (Clifton, NJ).

Acyl carrier protein (ACP) and acyl-ACP synthase are isolated from *E. coli* strain K-12 as described by Rock and Cronan (Rock and Cronan, *Methods in Enzymol* (1981) *71*:341-351 and Rock *et al.*, *Methods in Enzymol.* (1981) *72*:397-403). The *E. coli* is obtainable from Grain Processing (Iowa) as frozen late-logarithmic phase cells.

[9,10(n)-$^3$H]stearic acid is synthesized by reduction of [9,10(n)-$^3$H]oleic acid with hydrazine hydrate essentially as described by Johnson and Gurr (*Lipids* (1971) *6*:78-84). [9,10(n)-$^3$H]oleic acid (2 mCi), supplemented with 5.58mg unlabeled oleic acid to give a final specific radioactivity of 100mCi/mmol, is dissolved in 2ml of acetonitrile, acidified with 40μl of glacial acetic acid, and heated to 55°C. Reduction is initiated with 100μl of

60% (w/w) hydrazine hydrate; oxygen is bubbled through the mixture continuously. After each hour acetonitrile is added to bring the volume back to 2ml and an additional 100µl of hydrazine hydrate is added. At the end of 5 hr. the reaction is stopped by addition of 3ml of 2M HCl. The reaction products are extracted with three 3ml aliquots of petroleum ether and the combined ether extracts are washed with water, dried over sodium sulfate and evaporated to dryness. The dried reaction products are redissolved in 1.0ml acetonitrile and stored at -20°C. The distribution of fatty acid products in a 15µl aliquot is determined by preparation of phenacyl esters, which are then analyzed by thin layer chromatography on C-18 reverse phase plates developed with methanol:water:95:5 (v/v). Usually reduction to [9,10(n)-$^{3}$H]stearic acid is greater than 90%, a small amount of unreacted oleic acid may remain. The analysis is used to establish fraction of the total radioactivity that is present as stearate, and thereby to determine the exact substrate concentration in the enzyme assay.

Acyl-ACP substrates, including [9,10(n)-$^{3}$H] stearoyl-ACP are prepared and purified by the enzymatic synthesis procedure of Rock, Garwin, and Cronan (*Methods in Enzymol.* (1981) *72*:397-403).

Acyl carrier protein was covalently bound to Sepharose 4B by reaction of highly purified ACP with CNBr-activated Sepharose 4B as described by McKeon and Stumpf (*J. Biol. Chem.* (1982) *257*:12141-12147).

**Example 1**

In this example, an initial purification of *C. tinctorius* (safflower) desaturase, following the method of McKeon and Stumpf (*J. Biol. Chem.* (1982) *257*:12141-12142), is described.

*Assay:* In each of the following steps, the presence of the enzyme is detected radiometrically by measuring enzyme-catalyzed release of tritium from [9,10(n)-

$^{3}$H]stearoyl-ACP. Preparation of this substrate is described in "Materials" above.

The assay is performed by mixing 150µl water, 5ml dithiothreitol (100mM, freshly prepared in water), 10µl bovine serum albumin (10mg/ml in water), 15µl NADPH (25mM, freshly prepared in 0.1M Tricine-HCl, pH 8.2), 25µl spinach ferredoxin (2mg/ml Sigma Type III in water), 3µl NADPH:ferredoxin oxidoreductase (2.5 units/ml from Sigma), and 1 µl bovine liver catalase (800,000 units/ml from Sigma); after 10 min at room temperature, this mixture is added to a 13x100 mm screw-cap test tube containing 250µl sodium 1,4-piperazinediethanesulfonate (0.1M, pH 6.0). Finally, 10µl of the sample to be assayed is added and the reaction is started by adding 30µl of the substrate, [9,10(n)-$^{3}$H]stearoyl-ACP (100µCi/µmol, 10µM in 0.1M sodium 1,4-piperazinediethanesulfonate, pH 5.8). After sealing with a cap, the reaction is allowed to proceed for 10 min. while shaking at 23°C. The reaction is terminated by addition of 1.2ml of 5.8% tricholoracetic acid and the resulting precipitated acyl-ACP's are removed by centrifugation. The tritium released into the aqueous supernatant by the desaturase reaction is measured by liquid scintillation spectrometry. One unit of activity is defined as the amount of enzyme required to convert 1µmol of stearoyl-ACP to oleoyl-ACP, or to release 4µg-atoms of $^{3}$H per minute.

*Source tissue:* Developing *C. tinctorius* seeds from greenhouse grown plants are harvested between 16 and 18 days after flowering, frozen in liquid nitrogen and stored at -70°C until extracted.

*Acetone Powder:* Approximately 50g of frozen seeds are ground in liquid nitrogen and sieved to remove large seed coat pieces to provide a fine embryo powder. The powder is washed with acetone on a Buchner funnel until all yellow color is absent from the filtrate. The powder is then air dried and further processed as described below, or may be stored frozen for at least a year at -70°C without loss of enzyme activity.

*Acetone Powder Extract:* The dried acetone powder is weighed and triturated with ten times its weight of 20mM potassium phosphate, pH 6.8; the mixture is then centrifuged at 12,000 x g for 20 minutes and decanted through a layer of Miracloth (Calbiochem, La Jolla, CA).

*Ion Exchange Chromatography:* The acetone powder extract is then applied to a DEAE-cellulose column (Whatman DE-52) (1.5 x 12 cm) equilibrated with 20mM potassium phosphate, pH 6.8. The pass-through and a wash with one column-volume (20ml) of buffer are pooled.

*Affinity Chromatography:* An affinity matrix for purification of the desaturase is prepared by reacting highly purified *E. coli* ACP, with CNBr-activated Sepharose 4B (Sigma). ACP (120mg) is reduced by treatment with 1mM dithiothreitol for 30 min on ice, and then desalted on Sephadex G-10 (Pharmacia) equilibrated with 0.1M sodium bicarbonate, pH 6.0. The treated ACP (20 ml, 6 mg/ml) is then mixed with 20ml of CNBr-activated Sepharose 4B swollen in 0.1M sodium bicarbonate, pH 7.0, and the mixture is allowed to stand at 4°C for one day. The gel suspension is then centrifuged, washed once with 0.1M sodium bicarbonate, pH 7.0, and then treated with 40ml 0.1M glycine, pH 8.0, for 4 hours at room temperature to block unreacted sites. The gel is then washed for five cycles with alternating 50ml volumes of 0.5M NaCl in 0.1M sodium acetate, pH 4.0, and 0.5M NaCl in 0.1M sodium bicarbonate, pH 6.5, to remove non-covalently bound ligand. The gel is loaded into a column (1.5 x 11.2 cm) and equilibrated in 20mM potassium phosphate, pH 6.8.

The combined fractions from the DE-52 column are applied to the column, which is subsequently washed with one column volume (20ml) of the equilibration buffer, and then with 2.5 column volumes (50ml) of 300mM potassium phosphate, pH 6.8. Fractions are assayed for protein using the BCA Protein Assay Reagent (Pierce, Rockford, IL) to make sure that all extraneous protein has been eluted. Active Δ-9 desaturase is eluted from the column with 600mM potassium phosphate, pH 6.8. Active fractions are analyzed

by polyacrylamide gel electrophoresis in sodium dodecyl sulfate (SDS-PAGE) on 0.75mm thick 8 x 12 cm mini-gels according to the method of Laemmli (*Nature* (1970) *227*:680). The running gel contains 10% acrylamide in a 30/0.8 ratio of acrylamide to cross-linker bis-acrylamide. Those fractions containing a predominant band at approximately 43 kD are pooled and stored frozen at -70°C until final purification. The yield from 50g of seed tissue is is approximately 60µg of protein as measured by amino acid analysis.

Further purification as described in Example 2 or Example 3 is then applied to the fractions pooled from the ACP-Sepharose column separation.

**Example 2**

In this example, a protocol for the final purification of *C. tinctorius* desaturase is described. Seeds are treated in accordance with Example 1.

*Reverse-Phase HPLC:* Fractions from the ACP-Sepharose column are pooled and applied to a Vydac C4 reverse-phase column (0.45 x 15 cm) equilibrated in 0.1% TFA, 7% acetonitrile. After a 10 min wash with 0.1% TFA, the column is eluted with a gradient of increasing acetonitrile (7%-70% v/v) in 0.1% TFA over a period of 45 min. The flow rate is 0.5ml/min throughout. Eluting components are monitored by absorbance at 214 nm. Δ-9 desaturase elutes at about 42 min. (approximately 50% acetonitrile); the major contaminant protein remaining from ACP-affinity chromatography elutes at about 28 min. (approximately 30% acetonitrile). The substantially homogeneous desaturase, which is no longer active, is identified by SDS-PAGE, in which it exhibits a single band corresponding to a molecular weight of approximately 43 kD. The quantity of desaturase protein in the sample may be determined by amino acid analysis.

**Example 3**

In this example, a protocol for the final purification of *C. tinctorius* desaturase is described. Seeds are treated in accordance with Example 1.

*Reduction and Alkylation:* Protein is precipitated out of the pooled fraction solutions recovered from the ACP-Sepharose column with 10% (w/v) trichloroacetic acid, washed with cold (-20°C) acetone, and resuspended in 1 ml 500mM Tris-HCl, pH 8.6, containing 6M guanidine-HCl, 10mM EDTA, and 3.2 mM dithiothreitol. After 2 hours, 3.52 µmol [$^3$H]-iodoacetic acid (64µCi/µmol, New England Nuclear) is added, and the reaction is allowed to proceed at room temperature in the dark for 2 hours, at which time the reaction is terminated by addition of 1µl (15µmol) ß-mercaptoethanol. The sample is then re-precipitated with 10% (w/v) trichloroacetic acid, and the pellet again washed with cold (-20°C) acetone and resuspended in Laemmli's SDS-sample buffer (*Nature* (1970) *227*:680).

*SDS-Polyacrylamide Gel Electrophoresis:* The resulting sample is boiled for 5 min. and then applied to a 1.5 mm thick, 8 x 12 cm, SDS-polyacrylamide mini-gel prepared as described by Laemmli, *supra*. The running gel contains 17.5% acrylamide in a 30:0.13 ratio of acrylamide to cross-linking bis-acrylamide. Separation is achieved by electrophoresis at 15 mA, for 2 hours at 4°C.

*Blotting from SDS-gels to PVDF Membrane:* Proteins are recovered from the gel by electroblotting at 5 mA/cm$^2$ to a four-layer sandwich of polyvinylidenedifluoride (PVDF) membrane for 2 h at 4°C in a buffer containing 10mM CAPS ("3-[cyclohexylamino]-1-propane-sulfonic acid"), pH 11. The membranes must be wetted in 50% methanol, prior to exposure to the blotting buffer. After blotting, the membrane layers are stained for 1-2 min. in 0.02% Coomassie Blue in 50% methanol, and then destained in 50% methanol. The desaturase is identified as a band corresponding to a molecular weight of about 43 kD; the major contaminant runs at or near the dye front of the gel corresponding to a molecular weight less than 20 kD.

The desaturase band on the PVDF membrane may be applied directly to the Edman sequencer (Applied Biosystems Model 477A) for determination of the N-terminal sequence of the intact protein, or for more extensive sequence determination, may be eluted from the membrane in 40% acetonitrile to recover pure desaturase in solution. Acetonitrile is removed from the eluted desaturase by evaporation on a Speed-Vac (Savant; Farmingdale, NY), and the substantially homogeneous Δ-9 desaturase is resuspended in an appropriate buffer for subsequent proteolytic digestion as described in Example 4. The quantity of desaturase protein present may be determined by amino acid analysis.

Alternatively, if the sample is to be digested with trypsin or gluC protease to generate peptides for amino acid sequence analysis, proteins may be electroblotted to nitrocellulose membranes and stained with Ponceau or amido black.

**Example 4**

In this example, a method for the determination of the amino acid sequence of a desaturase is described.

*Reduction and Alkylation:* Substantially homogenous stearoyl-ACP desaturase (See, Example 2) is reduced and alkylated with [$^3$H]-iodacetic acid (See, Example 3), except that the final acetone-washed pellet is resuspended in the appropriate buffer for subsequent proteolysis. Reduction and alkylation assures complete denaturation of the protein so that complete proteolysis can occur. The sample may be alkylated with radiolabeled iodoacetamide or with 4-vinylpyridine instead of [$^3$H]-iodacetic acid in substantially the same manner. Use of iodoacetic acid affords an alkylated sample with greater solubility, which is advantageous in subsequent sample manipulation.

*Proteolysis:* Substantially pure alkylated samples are digested with endoproteinase lysC. The sample is resuspended in 100 µl of 25 mM Tris-HCl, pH 8.8, containing 1 mM EDTA. Endoproteinase lysC is added to the sample in a

protease/desaturase ratio of 1/50 (w/w). Digestion is allowed to proceed at room temperature for 8 hours, at which time another equal amount of protease is added. After 18 more hours, 1 µl of concentrated HCl is added to stop proteolysis, and the sample is applied directly to a Vydac C18 reverse-phase column (0.2 x 15 cm) equilibrated in 7% acetonitrile (v/v) in 0.1 mM sodium phosphate, pH 2.2. After washing for 20 min with the equilibration buffer, peptides are eluted with a gradient in acetonitrile (7-70%, v/v) over 120 min. Flow rate is 50 µl/min, throughout. Eluting components are monitored by absorbance at 214 nm, and individual peptide peaks are collected as separate fractions. The peptide fractions are further purified by application to a second Vydac C18 reverse-phase column (0.2 x 15 cm) equilibrated in 7% (v/v) acetonitrile in 0.1% (v/v) trifluoroacetic acid. Again, after a 20 min wash with equilibration buffer, the substantially pure peptides are eluted with a gradient (7-70%, v/v) of acetonitrile in 0.1% trifluoroacetic acid over 120 min. The flow rate is 50 µl/min, throughout. Eluting components are monitored by absorbance at 214 nm, and individual peptide peaks are collected as separate fractions. These substantially pure peptides are applied directly to the Edman sequencer (Applied Biosystems, Model 477A) for amino acid sequence analysis. Alternatively, peptide fraction from the first HPLC purification in phosphate buffer, or from a single chromatography step in trifluoroacetic acid buffer, may be applied directly to the sequencer, but these fractions, in many cases, are not substantially pure and yield mixed or ambiguous sequence information.

Other proteases may be used to digest desaturase, including but not limited to trypsin, gluC, and aspN. While the individual digest buffer conditions may be different, the protocols for digestion, peptide separation, purification, and sequencing are substantially the same as those outlined for the digestion with lysC. Alternatively, desaturase may be digested chemically using cyanogen bromide (Gross *Methods Enzymol* (1967) *11*:238-255 or Gross

and Witkop *J. Am. Chem. Soc.* (1961) *83*:1510), hydroxylamine (Bornstein and Balian *Methods Enzymol.* (1977) *47*:132-745), iodosobenzoic acid (Inglis *Methods Enzymol.* (1983) *91*:324-332), or mild acid (Fontana *et al.*, *Methods Enzymol.* (1983) *91*:311-317), as described in the respective references.

Fragments generated from these digestion steps of *C. tinctorius* desaturase are presented in Fig. 1 and as SEQ ID NOS: 1-11.

**Example 5**

In this example, the preparation of a plant embryo cDNA bank, using the methods as described in Alexander, *et al.* (*Methods in Enzymology* (1987) *154*:41-64) and the screening of the bank to obtain a desaturase cDNA clone is described.

*C. tinctorius*: A plant embryo cDNA library may be constructed from poly(A)+ RNA isolated from *C. tinctorius* embryos collected at 14-17 days post-anthesis. Poly(A)+ RNA is isolated from polyribosomes by a method initially described by Jackson and Larkins (*Plant Physiol.* (1976) *57*:5-10) as modified by Goldberg *et al.* (*Developmental Biol.* (1981) *83*:201-217).

The plasmid cloning vector pCGN1703, derived from the commercial cloning vector Bluescribe M13- (Stratagene Cloning Systems; San Diego, CA), is made as follows. The polylinker of Bluescribe M13- is altered by digestion with *Bam*HI, treatment with mung bean endonuclease, and blunt-end ligation to create a *Bam*HI-deleted plasmid, pCGN1700. pCGN1700 is digested with *Eco*RI and *Sst*I (adjacent restriction sites) and annealed with synthetic complementary oligonucleotides having the sequences 5' CGGATCCACTGCAGTCTAGAGGGCCCGGGA 3'(SEQ ID NO: 30) and 5' AATTTCCCGGGCCCTCTAGACTGCAGTGGATCCGAGCT 3' (SEQ ID NO: 31). These sequences are inserted to eliminate the *Eco*RI site, move the *Bam*HI site onto the opposite side of the *Sst*I (also, sometimes referred to as "*Sac*I" herein) site found in Bluescribe, and to include new restriction sites *Pst*I, *Xba*I, *Apa*I, *Sma*I. The resulting plasmid pCGN1702, is

digested with *Hin*dIII and blunt-ended with Klenow enzyme; the linear DNA is partially digested with *Pvu*II and ligated with T4 DNA ligase in dilute solution. A transformant having the *lac* promoter region deleted is selected (pCGN1703) and is used as the plasmid cloning vector.

Briefly, the cloning method for cDNA synthesis is as follows. The plasmid cloning vector is digested with *Sst*I and homopolymer T-tails are generated on the resulting 3'-overhang sticky-ends using terminal deoxynucleotidyl transferase. The tailed plasmid is separated from undigested or un-tailed plasmid by oligo(dA)-cellulose chromatography. The resultant vector serves as the primer for synthesis of cDNA first strands covalently attached to either end of the vector plasmid. The cDNA-mRNA-vector complexes are treated with terminal transferase in the presence of deoxyguanosine triphosphate, generating G-tails at the ends of the cDNA strands. The extra cDNA-mRNA complex, adjacent to the *Bam*HI site, is removed by *Bam*HI digestion, leaving a cDNA-mRNA-vector complex with a *Bam*HI sticky-end at one end and a G-tail at the other. This complex is cyclized using the annealed synthetic cyclizing linker,

5'-
GATCCGCGGCCGCGAATTCGAGCTCCCCCCCCCC-3' and
3'-GCGCCGGCGCTTAAGCTCGA-5'

which has a *Bam*HI sticky-end and a C-tail end. Following ligation and repair the circular complexes are transformed into *E. coli* strain DH5α (BRL; Gaithersburg, MD) to generate the cDNA library. The *C. tinctorius* embryo cDNA bank contains between $3x10^6$ and $5x10^6$ clones with an average cDNA insert size of approximately 1000 base pairs.

*Probe production Including PCR Reactions*: Two regions of amino acid sequence (Example 4) with low codon degeneracy are chosen from opposite ends of peptide sequence "Fragment F2" (SEQ ID NO:2) for production of a probe for the plant desaturase cDNA. Two sets of mixed oligonucleotides are designed and synthesized for use as forward (SEQ ID NOS: 21-24) and reverse (SEQ ID NOS: 25-26) primers, respectively, in the polymerase chain reaction

(Saiki *et al.*, *Science* (1985) *230*:1350-1354; Oste, *Biotechniques* (1988) *6*:162-167). See, Fig. 6. All oligonucleotides are synthesized on an Applied Biosystems 380A DNA synthesizer.

Probes to *C. tinctorius* desaturase may be prepared using the peptide sequence "Fragment 2" (SEQ ID NO: 2) identified in Fig. 1. Four types of forward primers were synthesized and labeled 13-1, 13-2, 13-3, and 13-4 (SEQ ID NOS: 21-24, respectively). Two groups of reverse primers were synthesized and designated 13-5A and 13-6A (SEQ ID NOS: 25-26, respectively). The primer sequences are shown in Fig. 6. These oligonucleotide groups have a redundancy of 64 or less and contain either 20 or 17 bases of coding sequence along with flanking restriction site sequences for *Hin*dIII or *Eco*RI. Based on the intervening amino acid sequence between the primer regions on peptide "Fragment 2" (SEQ ID NO: 2) the PCR product is expected to contain 107 base pairs.

Polymerase chain reaction is performed using the cDNA library DNA as template and the possible eight combinations of the four forward and two reverse oligonucleotides as primers in a Perkin-Elmer/Cetus DNA Thermal Cycler (Norwalk, CT) thermocycle file 1 min. 94°C, 2 min. 42°C, 2 min rise from 42°-72°C for 30 cycles, followed by the step cycle file without step rises, 1 min. 94°C, 2 min. 42°C, 3 min. 72°C with increasing 15 sec extensions of the 72°C step for 10 cycles, and a final 10 min. 72°C extension.

The product of the 13-4 forward primer (SEQ ID NO: 24) and the 13-5A reverse primer (SEQ ID NO: 25) reaction was ethanol precipitated and then digested with *Hin*dIII and *Eco*RI, the resulting fragment was subcloned into pUC8 (Vieira and Messing, *Gene* (1982) *19*:259-268). Minipreparation DNA (Maniatis *et al.*, *Molecular_Cloning: A Laboratory Manual* (1982) Cold Harbor Laboratory, New York) of one clone was sequenced by Sanger dideoxy sequencing (Sanger *et al.*, *Proc. Nat. Acad. Sci. USA* (1977) *74*:5463-5467) using the M13 universal and reverse primers. Translation of the resulting DNA sequence results in a

peptide sequence that exactly matches the amino acid sequence in peptide "Fragment F2" (SEQ ID NO: 2).

An exact 50 base oligonucleotide designated DESAT-50 is synthesized to match the sequence of the PCR reaction product from the first valine residue to the last tyrosine residue.

The probe DSAT-50 5' - GTAAGTAGGTAGGGCTTCCTCTGTAATCATATCTCCAACCAAAACAACAA -3' (SEQ ID NO: 32) is used to probe the *C. tinctorius* embryo cDNA library.

*Library screen*

The *C. tinctorius* embryo cDNA bank is moved into the cloning vector lambda gt10 (Stratagene Cloning Systems) by digestion of total cDNA with *Eco*RI and ligation to lambda gt10 DNA digested with *Eco*RI. The titer of the resulting library was ~$5x10^5$/ml. The library is then plated on *E. coli* strain C600 (Huynh, *et al.*, *DNA Cloning Vol. 1* Eds. Glover D.M. IRL Press Limited: Oxford England, pp. 56, 110) at a density of 5000 plaques/150 mm NZY ("NZYM" as defined in Maniatis *et al. supra*) agar plate to provide over 45,000 plaques for screening. Duplicate lifts are taken of the plaques using NEN Colony Plaque Screen filters by laying precut filters over the plates for ~1 minute and then peeling them off. The phage DNA is immobilized by floating the filters on denaturing solution (1.5M NaCl, .05M NaOH) for 1 min., transferring the filters to neutralizing solution (1.5M NaCl, 0.5M Tris-HCI pH 8.0) for 2 min. and then to 2XSSC (1xSSC = 0.15M NaCl; 0.015M Na citrate) for 3 min., followed by air drying. The filters are hybridized with $^{32}P$ end-labeled DSAT-50 oligonucleotide (SEQ ID NO: 32) (BRL 5' DNA Terminus Labeling System) by the method of Devlin et al., (*DNA* (1988) 7:499-807) at 42° C overnight, and washed for 30 min. at 50°C in 2XSSC, 0.5% SDS and then twice for 20 min. each at 50°C in 0.1XSSC, 0.5% SDS. Filters are exposed to X-ray film at -70°C with a Dupont Cronex intensifying screen for 48 hours.

Clones are detected by hybridization with the DSAT-50 oligonucleotide and plaque purified. The complete nucleotide sequence (SEQ ID NO: 12) of the cDNA insert of a clone, pCGN2754, and a partial restriction map thereof are presented in Figures 2 and 7A, respectively. The cDNA insert includes 1533 bases plus a poly(A) track at the 3' end of 100-200 bases. The open reading frame for the desaturase begins at the first ATG (nucleotide 106) from the 5' end and stops at nucleotide 1294. The translated amino acid sequence is presented in Fig. 2 and SEQ ID NO: 13. The open reading frame includes a 33 amino acid transit peptide not found in the amino acid sequence of the mature protein. The N-terminus of the protein begins at the alanine immediately following the *Nco*I site (nucleotide 202) indicating the site of the transit peptide processing.

**Example 6**

In this example, expression of a plant desaturase in a prokaryote is described.

*Desaturase expression construct in E. coli*

A plasmid for expression of desaturase activity in *E. coli* is constructed as follows. The desaturase cDNA clone pCGN2754 is digested with *Hin*dIII and SalI and ligated to pCGN2016 (a chloramphenicol resistant version of Bluescript KS-) digested with *Hin*dIII and *Xho*I. The resulting plasmid is pCGN1894.

pCGN2016 is prepared by digesting pCGN565 with *Hha*I, and the fragment containing the chloramphenicol resistance gene is excised, blunted by use of mung bean nuclease, and inserted into the *Eco*RV site of Bluescript KS- (Stratagene: La Jolla, CA) to create pCGN2008. The choramphenicol resistance gene of pCGN2008 is removed by *Eco*RI/*Hin*dIII digestion. After treatment with Klenow enzyme to blunt the ends, the fragment is ligated to *Dra*I digested Bluescript KS-. A clone that has the *Dra*I fragment containing ampicillin resistance replaced with the chloramphenicol resistance is chosen and named pCGN2016.

pCGN565 is a cloning vector based on pUC12-cm (K. Buckley Ph.D. Thesis, Regulation and expression of the phi X174 lysis gene, University of California, San Diego, 1985), but contains pUC18 linkers (Yanisch-Perron, *et al.*, *Gene* (1985) *53*:103-119).

The fragment containing the mature coding region of the Δ-9 desaturase, 3'-noncoding sequences and poly(A) tails is subcloned from pCGN1894 digested with *Nco*1 and *Asp*718 into pUC120, an *E. coli* expression vector based on pUC118 (Vieira and Messing, *Methods in Enzymology* (1987) *153*:3-11) with the lac region inserted in the opposite orientation and an *Nco*I site at the ATG of the lac peptide (Vieira, J. PhD. Thesis, University of Minnesota, 1988). The *E. coli* desaturase expression plasmid is designated pCGN3201. The desaturase sequences are inserted such that they are aligned with the lac transcription and translation signals.

*Expression of Desaturase in E.coli*

Single colonies of *E. coli* strain 7118 (Maniatis *et al.*, *supra*) containing pUC120 or pCGN3201 are cultured in 80 mls each of ECLB broth, 300 mg/L penicillin. The cells are induced by the addition of 1mM IPTG. Cells are grown overnight (18 hrs) at 37° C.

Eighty mls of overnight cultures of *E. coli* (induced and uninduced) containing pUC120 or pCGN3201 are centrifuged at 14,800 x g for 15 min. The pelleted cells are resuspended in 3 mls 20 mM phosphate buffer, pH 6.8. Resuspended cells were broken in a french press at 16,000 psi. Broken cell mixtures are centrifuged 5000xg for 5 min. 100 µl of each supernatant is applied to a G-25 Sephadex gel filtration centrifugal column (Boehringer Mannheim Biochemicals), equilibrated in 20mM phosphate buffer pH 6.8. Columns are spun for 4 min at 5000xg. Effluent was collected and used as enzyme source in the desaturase assay.

Desaturase activity is assayed as described in Example 1. Both pUC120-containing, IPTG-induced cells and uninduced cells do not express detectable stearoyl-ACP

desaturase activity. The pCGN3201 IPTG-induced extract contains 8.22 nmol/min of desaturase activity. pCGN3201 uninduced extracts contains 6.45 nmol/min of activity. The pCGN3201 IPTG-induced extract shows 21.5% more activity than the uninduced pCGN3201 extract.

*Detection of induced protein in E. coli.*

Extracts of overnight cultures of *E. coli* strain 7118 (Maniatis *et al. supra* ) containing pCGN3201 or pUC120 grown in ECLB containing 300 mg/L penicillin induced with 1mM IPTG are prepared as follows. 1.5 ml of overnight culture grown shaking at 37°C are pelleted in Eppendorf tubes for 10 min at 10-13,000 µg. Pellets are resuspended in 150 ul SDS sample buffer (0.05M Tris-HCl, pH6.8, 1% SDS, 5% ß-mercaptoethanol, 10% glycerol and 0.005% bromophenol blue) and boiled for 10 min. 25 µl of each sample are electrophoresed on a 10% polyacrylamide gel (Laemmli, *Nature* (1970) *227*:680) at 25 mA for 5 hours. Gels are stained in 0.05% Coomassie Brilliant Blue, 25% isopropanol and 10% acetic acid and destained in 10% acetic acid and 10% isopropanol. A band is detected at a position just below the 43,000 MW protein marker (SDS PAGE standard, Low molecular weight, BioRad, Richmond CA) in the pCGN3201 extracts that is not present in the pUC120 extracts. This is the approximate molecular weight of mature desaturase protein.

*Requirement for Spinach Ferredoxin*

Stearoyl-ACP desaturase can also be expressed in *E. coli* by subcloning into the *E. coli* expression vector pET8c (Studier, *et al., Methods Enzymol.* (1990) *185*:60-89). The mature coding region (plus an extra Met codon) of the desaturase cDNA with accompanying 3'-sequences is inserted as an *Nco*1 - *Sma* 1 fragment into pET8c at the *Nco*1 and *Bam*H1 sites (after treatment of the *Bam*H1 site with Klenow fragment of DNA polymerase to create a blunt end) to create pCGN3208. The plasmid pCGN3208 is maintained in *E. coli* strain BL21(DE3) which contains the T7 polymerase gene

under the control of the isopropyl-b-D-thiogalactopyranoside (IPTG)-inducible *lac*UV5 promoter (Studier *et al.*, *supra*).

*E. coli* cells containing pCGN3208 are grown at 37°C to an $OD_{595}$ of ~0.7 in NZY broth containing 0.4% (w/v) glucose and 300 mg/liter penicillin, and then induced for 3 hours with 0.4 mM IPTG. Cells are pelleted from 1 ml of culture, dissolved in 125 µl of SDS sample buffer (10) and heated to 100°C for 10 min. Samples are analyzed by SDS polyacrylamide gel electrophoresis at 25 mA for 5 h. Gels are stained in 0.05% Coomassie Brilliant Blue, 25% (v/v) isopropanol and 10% (v/v) acetic acid. A band is detected at a position just below the 43,000 MW protein marker (SDS PAGE standard, Low Molecular Weight, BioRad, Richmond, CA) in the pCGN3208 extract that is not present in the pET8c extracts. This is the approximate molecular weight of mature desaturase protein.

For activity assays, cells are treated as described above and extracts are used as enzyme source in the stearoyl-ACP desaturase assay as described in Example 1. The extract from IPTG-induced pCGN3208 cells contains 8.61 nmol/min/mg protein of desaturase activity. The extract from pCGN3208 uninduced cells contains 1.41 nmol/min/mg protein of desaturase activity. The extract from pCGN3208 induced cells, thus has approximately 6-fold greater activity than the extract from uninduced pCGN3208 cells. Extracts from both induced and uninduced cells of pET8c do not contain detectable stearoyl-ACP desaturase activity.

Samples are also assayed as described in Example 1, but without the addition of spinach ferredoxin, to determine if the *E. coli* ferredoxin is an efficient electron donor for the desaturase reaction. Minimal activity is detected in *E. coli* extracts unless spinach ferredoxin is added exogenously.

**Example 7**

In this example, the preparation of an ACP expression cassette containing a plant desaturase in a binary vector suitable for plant transformation is described.

*ACP Expression Cassette*

An expression cassette utilizing 5'-upstream sequences and 3'-downstream sequences obtainable from *B. campestris* ACP gene can be constructed as follows.

A 1.45kb *Xho*I fragment of Bcg 4-4 (Fig. 9 and SEQ ID NO: 28) containing 5'-upstream sequences is subcloned into the cloning/sequencing vector Bluescript+ (Stratagene Cloning Systems, San Diego, CA). The resulting construct, pCGN1941, is digested with *Xho*I and ligated to a chloramphenicol resistant Bluescript M13+ vector, pCGN2015 digested with *Xho*I. pCGN2015 is prepared as described for pCGN2016 (See, Example 6) except that the *Eco*RI/*Hin*dIII "chloramphenicol" fragment isolated from pCGN2008 is ligated with the 2273 bp fragment of Bluescript $KS^+$ (Stratagene; LaJolla, CA) isolated after digestion with *Dra*I. This alters the antibiotic resistance of the plasmid from penicillin resistance to chloramphenicol resistance. The chloramphenicol resistant plasmid is pCGN1953.

3'-sequences of Bcg 4-4 are contained on an *Sst*I/BglII fragment cloned in the *Sst*I/*Bam*HI sites of M13 Bluescript+ vector. This plasmid is named pCGN1940. pCGN1940 is modified by *in vitro* site-directed mutagenesis (Adelman *et al.*, *DNA* (1983) *2*:183-193) using the synthetic oligonucleotide 5'-CTTAAGAAGTAACCCGGGCTGCAGTTTTAGTATTAAGAG-3' (SEQ ID NO: 33) to insert *Sma*I and *Pst*I restriction sites immediately following the stop codon of the reading frame for the ACP gene 18 nucleotides from the *Sst*I site. The 3'-noncoding sequences from this modified plasmid, pCGN1950, are moved as a *PsI-Sma*I fragment into pCGN1953 cut with *Pst*I and *Sma*I. The resulting plasmid pCGN1977 comprises the ACP expression cassette with the unique restriction sites *Eco*RV, *Eco*RI and *Pst*I available between the 1.45kb 5' and 1.5 kb of 3'-noncoding sequences (SEQ ID

NO: 28) for the cloning of genes to be expressed under regulation of these ACP gene regions.

*Desaturase Expression in Plants*

Desaturase cDNA sequences from pCGN2754 are inserted in the ACP expression cassette, pCGN1977, as follows. pCGN2754 is digested with *Hin*dIII (located 160 nucleotides upstream of the start codon) and *Asp*718 located in the polylinker outside the poly(A) tails. The fragment containing the coding region for desaturase was blunt-ended using DNA polymerase I and ligated to pCGN1977 digested with *Eco*RV. A clone containing the desaturase sequences in the sense orientation with respect to the ACP promoter is selected and called pCGN1895. This expression cassette may be inserted into a binary vector, for example, for *Agrobacterium*-mediated transformation, or employed in other plant transformation techniques.

*Binary Vector and Agrobacterium Transformation*

The fragment containing the pCGN1895 expression sequences ACP 5'/desaturase/ACP 3' is cloned into a binary vector pCGN1557 (described below) for *Agrobacterium* transformation by digestion with Asp718 and XbaI and ligation to pCGN1557 digested with Asp718 and XbaI. The resulting binary vector is called pCGN1898.

pCGN1898 is transformed into *Agrobacterium tumefaciens* strain EHA101 (Hood, *et al.*, *J. Bacteriol.* (1986) *168*:1291-1301) as per the method of Holsters, *et al.*, *Mol. Gen. Genet.* (1978) *163*:181-187.

RNA blot analysis of seeds (T2) from T1 plants show the presence of a mRNA species for the inserted *C. tinctorius* desaturase, but the amount of message is low compared to endogenous levels of mRNA for the *Brassica* desaturase, suggesting that the expression levels were insufficient to significantly increase the amount of desaturase enzyme above that normally present. This is consistent with the negative results from oil, desaturase activity and Western blot analyses.

Construction of pCGN1557

pCGN1557 (McBride and Summerfelt, *Plant Molecular Biology* (1990) *14(2)*:269-276) is a binary plant transformation vector containing the left and right T-DNA borders of *Agrobacterium tumefaciens* octopine Ti-plasmid pTiA6 (Currier and Nester, *supra*, the gentamycin resistance gene of pPH1JI (Hirsch and Beringer, *supra*), an *Agrobacterium rhizogenes* Ri plasmid origin of replication from pLJbB11 (Jouanin *et al.*, *supra*), a 35S promoter-kanR-tml3' region capable of conferring kanamycin resistance to transformed plants, a ColE1 origin of replication from pBR322 (Bolivar *et al.*, *supra*), and a *lacZ'* screenable marker gene from pUC18 (Yanish-Perron *et al.*, *supra*).

There are three major intermediate constructs used to generate pCGN1557:

pCGN1532 (see below) contains the pCGN1557 backbone, the pRi plasmid origin of replication, and the ColE1 origin of replication.

pCGN1546 (see below) contains the CaMV35S5'-kan$^R$-tml3' plant selectable marker region.

pCGN1541b (see below) contains the right and left T-DNA borders of the *A. tumefaciens* octopine Ti-plasmid and the *lacZ'* region from pUC19.

To construct pCGN1557 from the above plasmids, pCGN1546 is digested with *Xho*I, and the fragment containing the CaMV 35S5'-kan$^R$-tml3' region is cloned into the *Xho*I site of pCGN1541b to give the plasmid pCGN1553, which contains T-DNA/left border/CaMV 35S5'-kanR-tml3'/*lacZ'*/T-DNA left border. pCGN1553 is digested with *Bgl*II, and the fragment containing the T-DNA/left border/CaMV35S5'-kan$^R$-tml3'/*lacZ'*/T-DNA left border region is ligated into *Bam*HI-digested pCGN1532 to give the complete binary vector, pCGN1557.

pCGN1532

The 3.5kb *Eco*RI-*Pst*I fragment containing the gentamycin resistance gene is removed from pPh1JI (Hirsch and Beringer, *Plasmid* (1984) *12*:139-141) by *Eco*RI-*Pst*I

digestion and cloned into *Eco*RI-*Pst*I digested pUC9 (Vieira and Messing, *Gene* (1982) *19*:259-268) to generate pCGN549. *Hin*dIII-*Pst*I digestion of pCGN549 yields a 3.1 kb fragment bearing the gentamycin resistance gene, which is made blunt ended by the Klenow fragment of DNA polymerase I and cloned into *Pvu*II digested pBR322 (Bolivar *et al.*, *Gene* (1977) *2*:95-113) to create pBR322Gm. pBR322Gm is digested with *Dra*I and *Sph*I, treated with Klenow enzyme to create blunt ends, and the 2.8 kb fragment cloned into the Ri origin-containing plasmid pLJbB11 (Jouanin *et al.*, *Mol. Gen. Genet.* (1985) *201*:370-374) which has been digested with *Apa*I and made blunt-ended with Klenow enzyme, creating pLHbB11Gm. The extra ColE1 origin and the kanamycin resistance gene are deleted from pLHbB11Gm by digestion with *Bam*HI followed by self closure to create pGmB11. The *Hin*dII site of pGmB11 is deleted by *Hin*dIII digestion followed by treatment with Klenow enzyme and self closure, creating pGmB11-H. The *Pst*I site of pGmB11-H is deleted by *Pst*I digestion followed by treatment with Klenow enzyme and self closure, creating pCGN1532.

Construction of pCGN1546

The 35S promoter-tml3' expression cassette, pCGN986, contains a cauliflower mosaic virus 35S (CaMV35) promoter and a T-DNA tml 3'-region with multiple restriction sites between them. pCGN986 is derived from another cassette, pCGN206, containing a CaMV35S promoter and a different 3' region, the CaMV region VI 3'-end. The CaMV 35S promoter is cloned as an *Alu*I fragment (bp 7144-7734) (Gardner *et. al.*, *Nucl. Acids Res.* (1981) *9*:2871-2888) into the *Hin*cII site of M13mp7 (Messing, *et. al.*, *Nucl. Acids Res.* (1981) *9*:309-321) to create C614. An *Eco*RI digest of C614 produced the *Eco*RI fragment from C614 containing the 35S promoter which is cloned into the *Eco*RI site of pUC8 (Vieira and Messing, *Gene* (1982) *19*:259) to produce pCGN147.

pCGN148a containing a promoter region, selectable marker (KAN with 2 ATG's) and 3' region, is prepared by

digesting pCGN528 with *Bgl*II and inserting the *Bam*HI-*Bgl*II promoter fragment from pCGN147. This fragment is cloned into the *Bgl*II site of pCGN528 so that the *Bgl*II site is proximal to the kanamycin gene of pCGN528.

The shuttle vector used for this construct pCGN528, is made as follows: pCGN525 is made by digesting a plasmid containing Tn5 which harbors a kanamycin gene (Jorgenson *et. al., Mol. Gen. Genet.* (1979) *177*:65) with *Hin*dIII-*Bam*HI and inserting the *Hin*dIII-*Bam*HI fragment containing the kanamycin gene into the *Hin*dIII-*Bam*HI sites in the tetracycline gene of pACYC184 (Chang and Cohen, *J. Bacteriol.* (1978) *134*:1141-1156). pCGN526 was made by inserting the *Bam*HI fragment 19 of pTiA6 (Thomashow *et. al., Cell* (1980) *19*:729-739), modified with *Xho*I linkers inserted into the *Sma*I site, into the *Bam*HI site of pCGN525. pCGN528 is obtained by deleting the small *Xho*I fragment from pCGN526 by digesting with *Xho*I and religating.

pCGN149a is made by cloning the *Bam*HI-kanamycin gene fragment from pMB9KanXXI into the *Bam*HI site of pCGN148a. pMB9KanXXI is a pUC4K variant (Vieira and Messing, *Gene* (1982) *19*:259-268) which has the *Xho*I site missing, but contains a functional kanamycin gene from Tn903 to allow for efficient selection in *Agrobacterium*.

pCGN149a is digested with *Hin*dIII and *Bam*HI and ligated to pUC8 digested with *Hin*dIII and *Bam*HI to produce pCGN169. This removes the Tn903 kanamycin marker. pCGN565 (see pCGN2016 description) and pCGN169 are both digested with *Hin*dIII and *Pst*I and ligated to form pCGN203, a plasmid containing the CaMV 35S promoter and part of the 5'-end of the Tn5 kanamycin gene (up to the *Pst*I site, Jorgenson *et. al.*, (1979), *supra*). A 3'-regulatory region is added to pCGN203 from pCGN204 (an *Eco*RI fragment of CaMV (bp 408-6105) containing the region VI 3' cloned into pUC18 (Yanisch-Perron, *et al., Gene* (1985) *33*:103-119) by digestion with *Hin*dIII and *Pst*I and ligation. The resulting cassette, pCGN206, is the basis for the construction of pCGN986.

The pTiA6 T-DNA tml 3'-sequences are subcloned from the *Bam*19 T-DNA fragment (Thomashow *et al.*, (1980) *supra*) as a *Bam*HI-*Eco*RI fragment (nucleotides 9062 to 12,823, numbering as in Barker *et al.*, *Plant Mol. Biol.* (1982) *2*:335-350) and combined with the pACYC184 (Chang and Cohen (1978), *supra*) origin of replication as an *Eco*RI-*Hin*dIII fragment and a gentamycin resistance marker (from plasmid pLB41), obtained from D. Figurski) as a *Bam*HI-*Hin*dIII fragment to produce pCGN417.

The unique *Sma*I site of pCGN417 (nucleotide 11,207 of the *Bam*19 fragment) is changed to a *Sac*I site using linkers and the *Bam*HI-*Sac*I fragment is subcloned into pCGN565 to give pCGN971. The *Bam*HI site of pCGN971 is changed to an *Eco*RI site using linkers. The resulting *Eco*RI-*Sac*I fragment containing the tml 3' regulatory sequences is joined to pCGN206 by digestion with *Eco*RI and *Sac*I to give pCGN975. The small part of the Tn5 kanamycin resistance gene is deleted from the 3'-end of the CaMV 35S promoter by digestion with *Sal*I and *Bgl*II, blunting the ends and ligation with *Sal*I linkers. The final expression cassette pCGN986 contains the CaMV 35S promoter followed by two *Sal*I sites, an *Xba*I site, *Bam*HI, *Sma*I, *Kpn*I and the tml 3' region (nucleotides 11207-9023 of the T-DNA).

The 35S promoter-tml 3' expression cassette, pCGN986 is digested with *Hin*dIII. The ends are filled in with Klenow polymerase and *Xho*I linkers added. The resulting plasmid is called pCGN986X. The *Bam*HI-*Sac*I fragment of pBRX25 (see below) containing the nitrilase gene is inserted into *Bam*HI-*Sac*I digested pCGN986X yielding pBRX66.

Construction of pBRX25 is described in U.S. Letters Patent 4,810,648, which is hereby incorporated by reference. Briefly, the method is as follows: The nucleotide sequence of a 1212-bp *Pst*I-*Hin*cII DNA segment encoding the bromoxynil-specific nitrilase contains 65-bp of 5' untranslated nucleotides. To facilitate removal of a portion of these excess nucleotides, plasmid pBRX9 is digested with *Pst*I, and treated with nuclease *Bal*31. *Bam*HI linkers are added to the resulting ends. *Bam*HI-*Hin*cII

fragments containing a functional bromoxynil gene are cloned into the *Bam*HI-*Sma*I sites of pCGN565. The resulting plasmid, pBRX25, contains only 11 bp of 5' untranslated bacterial sequence.

pBRX66 is digested with *Pst*I and *Eco*RI, blunt ends generated by treatment with Klenow polymerase, and *Xho*I linkers added. The resulting plasmid pBRX68 now has a tml 3' region that is approximately 1.1kb. pBRX68 is digested with *Sal*I and *Sac*I, blunt ends generated by treatment with Klenow polymerase and *Eco*RI linkers added. The resulting plasmid, pCGN986XE is a 35S promoter - tml 3' expression cassette lacking the nitrilase gene.

The Tn5 kanamycin resistance gene is then inserted into pCGN986XE. The 1.0 kb *Eco*RI fragment of pCGN1536 (see pCGN1547 description) is ligated into pCGN986XE digested with *Eco*RI. A clone with the Tn5 kanamycin resistance gene in the correct orientation for transcription and translation is chosen and called pCGN1537b. The 35S promoter $Kan^R$-tml 3' region is then transferred to a chloramphenical resistant plasmid backbone. pCGN786, (a pUC-CAM based vector with the synthetic oligonucleotide 5' GGAATTCGTCGACAGATCTCTGCAGCTCGAGGGATCCAAGCTT 3' (SEQ ID NO: 34) containing the cloning sites *Eco*RI, *Sal*I, *Bgl*II, *Pst*I, *Xho*I, *Bam*HI, and *Hin*dIII inserted into pCGN566, pCGN566 contains the *Eco*HI-*Hin*dIII linker of pUC18 inserted into the *Eco*KI-*Hin*dIII sites of pUC13-cm (K. Buckler (1985) *supra*)) is digested with *Xho*I and the *Xho*I fragment of pCGN1537b containing the 35S promoter - $Kan^R$-tml 3' region is ligated in. The resulting clone is termed pCGN1546.

**pCGN1541b**

pCGN565RBα2X (see below) is digested with *Bgl*II and *Xho*I, and the 728bp fragment containing the T-DNA right border piece and the *lacZ'* gene is ligated with *Bgl*II-*Xho*I digested pCGN65ΔKX-S+K (see below), replacing the *Bgl*II-*Xho*I right border fragment of pCGN65ΔKX-S+K. The resulting plasmid, pCGN65α2X contains both T-DNA borders and the *lacZ'* gene. The *Cla*I fragment of pCGN65α2X is

replaced with an *Xho*I site by digesting with *Cla*I blunting the ends using the Klenow fragment, and ligating with *Xho*I linker DNA, resulting in plasmid pCGN65α2XX. pCGN65α2XX is digested with *Bgl*II and *Eco*RV, treated with the Klenow fragment of DNA polymerase I to create blunt ends, and ligated in the presence of *Bgl*II linker DNA, resulting in pCGN65α2XX'. pCGN65α2XX' is digested with *Bgl*II and ligated with *Bgl*II digested pCGN1538 (see below), resulting in pCGN1541a, which contains both plasmid backbones. pCGN1541a is digested with *Xho*I and religated. Ampicillin resistant, chlormaphenicol sensitive clones are chosen, which lack the pACYC184-derived backbone, creating pCGN1541b.

pCGN1538 is generated by digesting pBR322 with *Eco*RI and *Pvu*II, treating with Klenow to generate blunt ends, and ligating with *Bgl*II linkers. pCGN1538 is ampicillin resistant, tetracycline sensitive.

**pCGN65ΔKX-S+K**

pCGN501 is constructed by cloning a 1.85 kb *Eco*RI-*Xho*I fragment of pTiA6 (Currier and Nester, *J. Bact.* (1976) *126*:157-165) containing bases 13362-15208 (Barker *et al.*, *Plant Mo. Biol.* (1983) *2*:335-350) of the T-DNA (right border), into *Eco*RI-*Sal*I digested M13mp9 (Vieira and Messing, *Gene* (1982) *19*:259-268). pCGN502 is constructed by cloning a 1.6 kb *Hin*dIII-*Sma*I fragment of pTiA6, containing bases 602-2212 of the T-DNA (left border), into *Hin*dIII-*Sma*I digested M13mp9. pCGN501 and pCGN502 are both digested with *Eco*RI and *Hin*dIII and both T-DNA-containing fragments cloned together into *Hin*dIII digested pUC9 (Vieira and Messing, *Gene* (1982) *19*:259-268) to yield pCGN503, containing both T-DNA border fragments. pCGN503 is digested with *Hin*dIII and *Eco*RI and the two resulting *Hin*dIII-*Eco*RI fragments (containing the T-DNA borders) are cloned into *Eco*RI digested pHC79 (Hohn and Collins, *Gene* (1980) *11*:291-298) to generate pCGN518. The 1.6kb *Kpn*I-*Eco*RI fragment from pCGN518, containing the left T-DNA border, is cloned into *Kpn*I-*Eco*RI digested pCGN565 to

generate pCGN580. The *Bam*HII-*Bgl*II fragment of pCGN580 is cloned into the *Bam*HI site of pACYC184 (Chang and Cohen, *J. Bacteriol.* (1978) *134*:1141-1156) to create pCGN51. The 1.4 kb *Bam*HI-*Sph*I fragment of pCGN60 containing the T-DNA right border fragment, is cloned into *Bam*HI-*Sph*I digested pCGN51 to create pCGN65, which contains the right and left T-DNA borders.

pCGN65 is digested with *Kpn*I and *Xba*I, treated with Klenow enzyme to create blunt ends, and ligated in the presence of synthetic *Bgl*II linker DNA to create pCGN65ΔKX. pCGN65ΔKX is digested with *Sal*I, treated with Klenow enzyme to create blunt ends, and ligated in the presence of synthetic *Xho*I linker DNA to create pCGN65ΔKX-S+X.

pCGN565RBα2X

pCGN451 (see below) is digested with *Hpa*I and ligated in the presence of synthetic *Sph*I linker DNA to generate pCGN55. The *Xho*I-*Sph*I fragment of pCGN55 (bp13800-15208, including the right border, of *Agrobacterium tumefaciens* T-DNA; (Barker *et al.*, *Gene* (1977) *2*:95-113) is cloned into *Sal*I-*Sph*I digested pUC19 (Yanisch-Perron *et al.*, *Gene* (1985) *53*:103-119) to create pCGN60. The 1.4 kb *Hin*dIII-*Bam*HI fragment of pCGN60 is cloned into *Hin*dIII-*Bam*HI digested pSP64 (Promega, Inc.) to generate pCGN1039. pCGN1039 is digested with *Sma*I and *Nru*I (deleting bp14273-15208; (Barker *et al.*, *Gene* (1977) *2*:95-113) and ligated in the presence of synthetic *Bgl*II linker DNA creating pCGN1039ΔNS. The 0.47 kb *Eco*RI-*Hin*dIII fragment of pCGN1039ΔNS is cloned into *Eco*RI-*Hin*dIII digested pCGN565 to create pCGN565RB. The *Hin*dIII site of pCGN565RB is replaced with an *Xho*I site by digesting with *Hin*dIII, treating with Klenow enzyme, and ligating in the presence of synthetic *Xho*I linker DNA to create pCGN565RB-H+X.

pUC18 (Norrander *et al.*, *Gene* (1983) *26*:101-106) is digested with *Hae*II to release the *lacZ'* fragment, treated with Klenow enzyme to create blunt ends, and the *lacZ'*-containing fragment ligated into pCGN565RB-H+X, which had been digested with *Acc*I and *Sph*I and treated with Klenow

enzyme in such a orientation that the *lacZ'* promoter is proximal to the right border fragment; this construct, pCGN565RBα2x is positive for *lacZ'* expression when plated on an appropriate host and contains bp 13990-14273 of the right border fragment (Barker *et al., Plant Mo. Biol.* (1983) *2*:335-350) having deleted the *AccI-SphI* fragment (bp 13800-13990).

pCGN451

pCGN451 contains an ocs5'-ocs3' cassette, including the T-DNA right border, cloned into a derivative of pUC8 (Vieira and Messing, *supra*). The modified vector is derived by digesting pUC8 with *Hinc*II and ligating in the presence of synthetic linker DNA, creating pCGN416, and then deleting the *Eco*RI site of pCGN416 by *Eco*RI digestion followed by treatment with Klenow enzyme and self-ligation to create pCGN426.

The ocs5'-ocs3' cassette is created by a series of steps from DNA derived from the octopine Ti-plasmid pTiA6 (Currier and Nester, *supra*). To generate the 5'end, which includes the T-DNA right border, an *Eco*RI fragment of pTiA6 (bp 13362-16202 (the numbering is by Barker, *et al., (Plant Mol. Bio* (1983) *2*:335-350) for the closely related Ti plasmid pTil5955)) is removed from pVK232 (Knauf and Nester, *Plasmid* (1982) *8*:45) by *Eco*RI digestion and cloned into *Eco*RI digested pACYC184 (Chang and Cohen, *supra*) to generate pCGN15.

The 2.4kb *Bam*HI-*Eco*RI fragment (bp 13774-16202) of pCGN15 is cloned into *Eco*RI-*Bam*HI digested pBR322 (Bolivar, *et al., supra*) to yield pCGN429. The 412 bp *Eco*RI-*Bam*HI fragment (bp 13362-13772) of pCGN15 is cloned into *Eco*RI-*Bam*HI digested pBR322 to yield pCGN407. The cut-down promoter fragment is obtained by digesting pCGN407 with *Xmn*I (bp 13512), followed by resection with *Bal*31 exonuclease, ligation of synthetic *Eco*RI linkers, and digestion with *Bam*HI. Resulting fragments of approximately 130 bp are gel purified and cloned into M13mp9 (Vieira and Messing, *supra*) and sequenced. A clone, I-4, in which the *Eco*RI linker has been inserted at bp 1362 between the

transcription initiation point and the translation initiation codon is identified by comparison with the sequence of de Greve, *et al.*, (*J. Mol. Appl. Genet.* (1982) *1*:499-512). The *Eco*RI cleavage site is at position 13639, downstream from the mRNA start site. The 141 bp *Eco*RI-*Bam*HI fragment of I-4, containing the cut-down promoter, is cloned into *Eco*RI-*Bam*HI digested pBR322 to create pCGN428. The 141 bp *Eco*RI-*Bam*HI promoter piece from pCGN428, and the 2.5 kb *Eco*RI-*Bam*HI ocs5' piece from pCGN429 are cloned together into *Eco*RI digested pUC19 (Vieira and Messing, *supra*) to generate pCGN442, reconstructing the ocs upstream region with a cut-down promoter section.

To generate the ocs3' end, the *Hin*dIII fragment of pLB41 (D. Figurski, UC San Diego) containing the gentamycin resistance gene is cloned into *Hin*dIII digested pACYC184 (Chang and Cohen, *supra*) to create pCGN413b. The 4.7 kb *Bam*HI fragment of pTiA6 (*supra*), containing the ocs3' region, is cloned into *Bam*HI digested pBR325 (F. Bolivar, *Gene* (1978) *4*:121-136) to create 33c-19. The *Sma*I site at position 11207 (Barker, *supra*) of 33c-19 is converted to an *Xho*I site using a synthetic *Xho*I linker, generating pCCG401.2. The 3.8 kb *Bam*HI-*Eco*RI fragment of pCGN401.2 is cloned into *Bam*HI-*Eco*RI digested pCGN413b to create pCGN419.

The ocs5'-ocs3' cassette is generated by cloning the 2.64 kb *Eco*RI fragment of pCGN442, containing the 5' region, into *Eco*RI digested pCGN419 to create pCNG446. The 3.1kb *Xho*I fragment of pCGN446, having the ocs5' region (bp 13639-15208) and ocs3' region (bp 11207-12823), is cloned into the *Xho*I site of pCGN426 to create pCGN451.

**Example 8**

In this example, the preparation of a Bce-4 expression cassette containing a plant desaturase is described.

The desaturase cDNA clone from pCGN2754 prepared as described in Example 5, is modified by *in vitro* mutagenesis to insert restriction sites immediately upstream of the ATG start codon and downstream of the TGA stop codon. A

single-stranded template DNA is prepared for the mutagenesis reaction from pCGN1894 (described in Example 6) as described by Messing, (*Methods in Enzymol.* (1983) *101*:20-79). Synthetic oligonucleotides are synthesized on an Applied Biosystems 380A DNA synthesizer. The oligonucleotides used are

5'-CCATTTTTGATCTTCCTCGAGCCCGGGCTGCAGTTCTTCTTCTTCTTG-3' (SEQ ID NO: 35) for the 5'mutagenesis and

5'-GCTCGTTTTTTTTTTCTCTGCAGCCCGGGCTCGAGTCACAGCTTCACC -3' (SEQ ID NO: 36) for the 3'-mutagenesis; both add *Pst*I, *Sma*I and *Xho*I sites flanking the coding region. Both oligonucleotides are 5'-phosphorylated (BRL 5'-Terminus labelling kit) and used for mutagenesis with the pCGN1894 template by the procedure of Adelman *et al.* (*DNA* (1983) *2*:183-193). Alternatively, the desired restriction sites may be inserted by PCR, using the 3' oligo described above (SEQ ID NO: 36) and another oligo,

5' ACTGACTGCAGCCCGGGCTCGAGGAAGATCAAAAATGGCTCTTC 3' (SEQ ID NO: 37) for the 3' and 5' primers, respectively. The template in this polymerase chain reaction is DNA from pCGN1894. The *Xho*I fragment from the resulting clone can be subcloned into the Bce4 expression cassette, pCGN1870 (described below) at the unique *Xho*I site. This Bce4/desaturase expression cassette can then be inserted in a suitable binary vector, transformed into *Agrobacterium tumefaciens* strain EHA101 and used to transform plants as provided in Example 10.

Bce-4 Expression Cassette

pCGN1870 is a Bce-4 expression cassette containing 5' and 3' regulatory regions of the Bce-4 gene and may be derived from the Bce-4 sequence found in pCGN1857, which was deposited with the ATCC on March 9, 1990, and assigned accession number 68251, or by methods known to one skilled in the art from the sequence (SEQ ID NO: 27) provided in Fig. 8. The Bce 4 gene may be isolated as follows:

The *Cla*I fragment of pCGN1857, containing the Bce4 gene is ligated into *Cla*I digested Bluescript KS+

(Stratagene; La Jolla, CA), producing pCGN1864. Single stranded DNA is made from pCGN1864 and altered by *in vitro* mutagenesis using the oligonucleotides

BCE45P:

(5'GAGTAGTGAACTTCATGGATCCTCGAGGTCTTGAAAACCTAGA3') (SEQ ID NO: 38) and

BCE43P:

(5'CAATGTCTTGAGAGATCCCGGGATCCTTAACAACTAGGAAAAGG3') (SEQ ID NO: 39)

as described by Adelman *et al.* (*DNA* (1983) *2*:183-193). The oligonucleotide BSCP2 (5'GTAAGACACGACTTATCGCCACTG3') (SEQ ID NO: 40), complementary to a portion of Bluescript, is included in the reaction to improve the yield of double-stranded DNA molecules. The resulting plasmid, pCGN1866, contains *Xho*I and *Bam*HI sites (from BCE45P) immediately 5' to the Bce4 start codon and *Bam*HI and *Sma*I sites (from BCE43P) immediately 3' to the Bce4 stop codon. The *Cla*I fragment of pCGN1866, containing the mutagenized sequences, is inserted into the *Cla*I site of pCGN2016 (described in Example 6), producing pCGN1866C. The *Cla*I fragment of pCGN1866C is used to replace the corresponding wild-type *Cla*I fragment of PCGN1867 (described below) to produce pCGN1868. Bce4 coding sequences are removed by digestion of pCGN1868 with *Bam*HI and recircularization of the plasmid to produce pCGN1870. The Bce4 expression cassette, pCGN1870, contains 7.4 kb of 5' regulatory sequence and 1.9 kb of 3' regulatory sequence derived from the Bce4 genomic clone separated by the cloning sites, *Xho*I, *Bam*HI, and *Sma*I. Desaturase sequences in sense or anti-sense orientation may be inserted into the cassette via the cloning sites and the resulting construct may be employed in a plant transformation technique.

pCGN1867

The *Bam*HI and *Sma*I sites of pUC18 are removed by *Bam*HI-*Sma*I digestion and recircularizing of the plasmid, without repair of the ends, to produce pCGN1862 The *Pst*I fragment of pCGN1857, containing the Bce4 gene, is inserted into the *Pst*I site of pCGN1862 to produce pCGN1867.

**Example 9**

In this example, the preparation of a napin 1-2 expression cassette containing a plant desaturase is described.

*Preparation of Desaturase Clone*

The desaturase cDNA clone from pCGN2754 is prepared and modified as described in Example 8. The *Xho*I fragment from the resulting clone can be subcloned into the napin 1-2 expression cassette, pCGN1808 (described below) at the unique *Xho*I site. This napin 1-2/desaturase expression cassette can then be inserted into a suitable binary vector, transformed into *A. tumefaciens* strain EHA101 in a like manner as described in Example 7.

Alternatively, the desaturase safflower clone may be prepared such that restriction sites flank the translation start and stop sites, as described in Example 8, with the following modification. PCR was carried out according to manufacturer's instructions except for the initial annealing of the oligonucleotides to the template. The reaction mix was heated to 90°C for 5 min, cooled to 37°C over a one hour period, kept at 37°C for 20 min and then subjected to standard PCR cycles. The PCR product was digested with *Pst*I and ligated to pUC8 (Vieira and Messing (1982) *Gene 19*:2359-268) digested with *Pst*I to produce pCGN3220. The *Nco*I/*Sac*I fragment of pCGN3220 containing the pUC8 vector and the 5' and 3' sequences of the safflower desaturase cDNA was gel purified and ligated to the gel-purified cloned *Nco*I/*Sac*I fragment from pCGN1894 (see Example 6). The resulting plasmid pCGN3222 contains safflower desaturase cDNA sequences partially from the cDNA

clone and partially from the PCR. The regions obtained from the PCR were confirmed by DNA sequencing as being identical to the original cloned sequence.

*Expression Cassettes*

Napin 1-2 pCGN1808 Expression Cassette

An expression cassette utilizing 5' upstream sequences and 3' downstream sequences obtainable from *B. campestris* napin gene can be constructed as follows.

A 2.7 kb *Xho*I fragment of napin 1-2 (Fig. 10 and SEQ ID NO: 29) containing 5' upstream sequences is subcloned into pCGN789 (a pUC based vector the same as pUC119 with the normal polylinker replaced by the synthetic linker - 5'GGAATTCGTCGACAGATCTCTGCAGCTCGAGGGATCCAAGCTT 3', SEQ ID NO: 41, (which represented the polylinker *Eco*RI, *Sal*I, *Bgl*II, *Pst*I, *Xho*I, *Bam*HI, *Hin*dIII) and results in pCGN940. The majority of the napin coding region of pCGN940 was deleted by digestion with *Sal*I and religation to form pCGN1800. Single-stranded DNA from pCGN1800 was used in an *in vitro* mutagenesis reaction (Adelman *et al.*, *DNA* (1983) *2*:183-193) using the synthetic oligonucleotide 5' GCTTGTTCGCCATGGATATCTTCTGTATGTTC 3', SEQ ID NO: 42. This oligonucleotide inserted an *Eco*RV and an *Nco*I restriction site at the junction of the promoter region and the ATG start codon of the napin gene. An appropriate mutant was identified by hybridization to the oligonucleotide used for the mutagenesis and sequence analysis and named pCGN1801.

A 1.7 kb promoter fragment was subcloned from pCGN1801 by partial digestion with *Eco*RV and ligation to pCGN786 (a pCGN566 chloramphenicol based vector with the synthetic linker described above in place of the normal polylinker) cut with *Eco*RI and blunted by filling in with DNA Polymerase I Klenow fragment to create pCGN1802.

A 2.1 kb *Sal*I fragment of napin 1-2 (Fig. 10 and SEQ ID NO: 29) containing 3' downstream sequences is subcloned into pCGN789 (described above) and results in pCGN941. pCGN941 is digested with *Xho*I and *Hin*dIII and the resulting approximately 1.6 kb of napin 3' sequences are inserted

into *Xho*I-*Hin*dIII digested pCGN1802 to result in pCGN1803. In order to remove a 326 nucleotide *Hin*dIII fragment inserted opposite to its natural orientation, as a result of the fact that there are 2 *Hin*dIII sites in pCGN1803, the pCGN1803 is digested with *Hin*dIII and religated. Following religation, a clone is selected which now contains only 1.25 kb of the original 1.6 napin 3' sequence. This clone, pCGN1808 is the napin 1-2 expression cassette and contains 1.725 kb of napin promoter sequences and 1.265 kb of napin 3' sequence with the unique cloning sites *Sal*I, *Bgl*II, *Pst*I and *Xho*I in between.

Napin 1-2 pCGN3223 Expression Cassette

Alternatively, pCGN1808 may be modified to contain flanking restriction sites to allow movement of only the expression sequences and not the antibiotic resistance marker to binary vectors such as pCGN1557 (McBride and Summerfelt, *supra*). Synthetic oligonucleotides containing *Kpn*I, *Not*I and *Hin*dIII restriction sites are annealed and ligated at the unique *Hin*dIII site of pCGN1808, such that only one *Hin*dIII site is recovered. The resulting plasmid, pCGN3200 contains unique *Hin*dIII, *Not*I and *Kpn*I restriction sites at the 3'-end of the napin 3'-regulatory sequences as confirmed by sequence analysis.

The majority of the napin expression cassette is subcloned from pCGN3200 by digestion with *Hin*dIII and *Sac*I and ligation to *Hin*dIII and *Sac*I digested pIC19R (Marsh, *et al.* (1984) *Gene 32*:481-485) to make pCGN3212. The extreme 5'-sequences of the napin promoter region are reconstructed by PCR using pCGN3200 as a template and two primers flanking the *Sac*I site and the junction of the napin 5'-promoter and the pUC backbone of pCGN3200 from the pCGN1808 construct. The forward primer contains *Cla*I, *Hin*dIII, *Not*I, and *Kpn*I restiction sites as well as nucleotides 408-423 of the napin 5'-sequence (from the *Eco*RV site) and the reverse primer contains the complement to napin sequences 718-739 which include the unique *Sac*I site in the 5'-promoter. The PCR was performed using a Perkin Elmer/Cetus

thermocycler according to manufacturer's specifications. The PCR fragment is subcloned as a blunt-ended fragment into pUC8 (Vieira and Messing (1982) *Gene 19*:259-268) and digested with *Hinc*II to give pCGN3217. Sequence of pCGN3217 across the napin insert verifies that no improper nucleotides were introduced by PCR. The napin 5-sequences in pCGN3217 are ligated to the remainder of the napin expression cassette by digestion with *Cla*I and *Sac*I and ligation to pCGN3212 digested with *Cla*I and *Sac*I. The resulting expression cassette pCGN3221, is digested with *Hin*dIII and the napin expression sequences are gel purified away and ligated to pIC20H (Marsh, *supra*) digested with *Hin*dIII. The final expression cassette is pCGN3223, which contains in an ampicillin resistant background, essentially identical 1.725 napin 5' and 1.265 3' regulatory sequences as found in pCGN1808. The regulatory regions are flanked with *Hin*dIII, *Not*I and *Kpn*I restriction sites and unique *Sal*I, *Bgl*II, *Pst*I, and *Xho*I cloning sites are located between the 5' and 3' noncoding regions.

Desaturase sequences in sense or anti-sense orientation may be inserted into a napin expression cassette via the cloning sites. The resulting construct may be employed for plant transformation. For example, one of ordinary skill in the art could also use known techniques of gene cloning, mutations, insertion and repair to allow cloning of a napin expression cassette into any suitable binary vector, such as pCGN1557 (described in Example 7) or other similar vectors.

*Desaturase Expression*

The coding region of the safflower desaturase contained in pCGN3222 is cloned into the pCGN3223 napin cassette by digestion with *Xho*I and ligation to pCGN3223 digested with *Xho*I and *Sal*I. The resulting plasmid, pCGN3229 is digested with *Asp*718 and inserted in the binary vector pCGN1578 (McBride and Summerfelt (1990) *Plant Mol. Biol. 14*:269-276) at the unique *Asp*718 site. The resulting binary vector is pCGN3231 and contains the safflower

desaturase coding sequences flanked by the napin 5' and 3' regulatory sequences as well as the plant selectable marker construct, 35s/NPTII/tml.

The resulting binary vector, pCGN3231, is transformed into *Agrobacterium* and utilized for plant transformation as described in Example 10. For Northern analysis, total RNA is isolated from day 21 and day 28 post-anthesis developing seed from plants transformed with pCGN3231. Five samples were analyzed at day 21 and two at day 28 post-anthesis. RNA was isolated by the method of Hughes and Galau (*Plant Mol. Biol. Reporter* (1988) *6*: 253-257). Northern blot analysis was performed using a labeled 0.8 kb *Bgl*II fragment of pCGN1894 as a probe. Prehybridization and hybridization was at 42°C in 50% formamide, 10X Denhardt's solution, 5X SSC, 0.1% SDS, 5mM EDTA and 100ug/ml denatured salmon sperm DNA. Filters were washed at 55°C in 0.1 X SSC, 0.1% SDS. Under these conditions, the probe does not hybridize to the endogenous *Brassica* desaturase gene sequences. mRNA complementary to the safflower desaturase was detected in all the transgenic samples examined. More mRNA was present at day 28 than at day 21 post-anthesis and the highest level of RNA was seen in transgenic 3231-8. The total safflower desaturase mRNA level was estimated to be ~0.01% of the message at day 28 post-anthesis.

Western analysis (see below) gives a preliminary indication of increased protein in one transformant, 3231-8. However, the Western analysis is complicated by two factors: 1. The presence of cross-reacting material at the same molecular weight as expected for the safflower desaturase. We believe this material is the endogenous *Brassica* desaturase. 2. The analysis of levels of protein expressed is also complicated by the normal developmental increase in the expression of desaturase protein during this time period. If the seeds examined are not at the precise developmental stage as the control seeds, quantitative differences in the amount of material seen may be simply due to the normal increase in the *Brassica*

desaturase over this time period and not due to the expression of the safflower desaturase.

Western Analysis

Soluble protein is extracted from developing seeds of *Brassica* by homogenization with one volume (1ml/gram fresh weight) of buffer containing 20mM potassium phosphate, pH 6.8. The homogenate is clarified by centrifugation at 12,000 x *g* for 10 minutes. A second centrifugation is performed if necessary to provide a non-particulate supernatant.

Protein concentration of the extract is measured by the micromethod of Bradford (*Anal. Biochem.* (1976) *72*:248-254). Proteins (20-60µg) are separated by denaturing electrophoresis by the method of Laemmli (*supra*), and are transferred to nitrocellulose membrane by the method of Towbin *et al.* (*Proc. Nat. Acad Sci.* (1979) *76*:4350-4354).

The nitrocellulose membrane is blocked by incubation at room temperature for 15 minutes or at 4°C overnight in Tris-buffered saline with Tween 20 (Polyoxyethylenesorbitan monolaurate) and "TTBS-milk", (TTBS = 20mM Tris-HCl, 500mM NaCl, 0.1% Tween 20 (v/v), pH 7.5; "TTBS-milk" = TTBS and 3% skim milk powder). The volume of liquid in all incubations with the nitrocellulose membrane is sufficient to cover the membrane completely. The membrane is then incubated for an additional 5 minutes in TTBS.

The nitrocellulose membrane is incubated for at least one hour with shaking at room temperature with rabbit anti-stearoyl-ACP desaturase antiserum that was diluted 5,000- or 10,000-fold in "TTBS-milk". The rabbit anti-desaturase antiserum was commercially prepared from desaturase protein (purified as described in Example 1) by Berkeley Antibody Co. (Richmond, CA). The membrane is washed twice by shaking with TTBS for 5 minutes and then with deionized $H_20$ for 30 seconds.

The nitrocellulose membrane is incubated for at least 45 minutes at room temperature in a solution of "TTBS-milk" in which anti-rabbit IgG-alkaline phosphatase conjugate

(Promega, Madison, WI) is diluted 7,500-fold. The membrane is washed twice in TTBS followed by deionized $H_2O$, as described above.

The nitrocellulose membrane is equilibrated in buffer containing 100mM Tris-HCl, 100mM NaCl, 50mM $MgCl_2$, pH 9.5, by shaking for 5 minutes. The color reaction is initiated by placing the nitrocellulose membrane into 50ml of the same buffer to which has been added 15mg *p*-nitroblue tetrazolium chloride and 7.5mg 5-bromo- 4 chloro- 3-indolyl phosphate toluidine salt (BioRad Labs; Richmond, CA). The color reaction is stopped by rinsing the nitrocellulose membrane with deionized $H_2O$ and drying it between filter papers.

Oil analysis of developing seeds indicated no significant change in oil composition of the transformed plants with respect to the control plants. This result is consistant with the low levels of safflower mRNA observed in transgenic plants as compared to levels of endogenous *Brassica* desaturase (Example 12).

**Example 10**

In this example, an *Agrobacterium*-mediated plant transformation is described. *Brassica napus* is exemplified. The method is also useful for transformation of other *Brassica* species including *Brassica campestris*.

*Plant Material and Transformation*

Seeds of *Brassica napus* cv. Delta are soaked in 95% ethanol for 2 min, surface sterilized in a 1.0% solution of sodium hypochlorite containing a drop of Tween 20 for 45 min., and rinsed three times in sterile, distilled water. Seeds are then plated in Magenta boxes with 1/10th concentration of Murashige minimal organics medium (Gibco) supplemented with pyrodoxine (50 µg/l), nicotinic acid (50 µg/l), glycine (200 µg/l), and 0.6% Phytagar (Gibco) pH 5.8. Seeds are germinated in a culture room at 22°C in a 16 h photoperiod with cool fluorescent and red light of

intensity approximately 65 μEinsteins per square meter per second ($\mu Em^{-2}S^{-1}$).

Hypocotyls are excised from 7 day old seedlings, cut into pieces approximately 4 mm in length, and plated on feeder plates (Horsch et al. 1985). Feeder plates are prepared one day before use by plating 1.0 ml of a tobacco suspension culture onto a petri plate (100x25 mm) containing about 30 ml MS salt base (Carolina Biological) 100 mg/l inositol, 1.3 mg/l thiamine-HCl, 200 mg $KH_2PO_4$ with 3% sucrose, 2,4-D (1.0 mg/l), 0.6% Phytagar, and pH adjusted to 5.8 prior to autoclaving (MS0/1/0 medium). A sterile filter paper disc (Whatman 3 mm) is placed on top of the feeder layer prior to use. Tobacco suspension cultures are subcultured weekly by transfer of 10 ml of culture into 100 ml fresh MS medium as described for the feeder plates with 2,4-D (0.2 mg/l), Kinetin (0.1 mg/l). All hypocotyl explants are preincubated on feeder plates for 24 h. at 22°C in continuous light of intensity 30 $\mu Em^{-2}S^{-1}$ to 65 $\mu EM^{-2}S^{-1}$.

Single colonies of *A. tumefaciens* strain EHA101 containing a binary plasmid are transferred to 5 ml MG/L broth and grown overnight at 30°C. Per liter, MG/L broth contains 5g mannitol, 1 g L-glutamic acid or 1.15 g sodium glutamate, 0.25 g $kH_2PO_4$, 0.10 g NaCL, 0.10 g $MGSO_4 \cdot 7H_20$, 1 mg biotin, 5 g tryptone, and 2.5 g yeast extract, and the broth is adjusted to pH 7.0. Hypocotyl explants are immersed in 7-12 ml MG/L broth with bacteria diluted to $1x10^8$ bacteria/ml and after 10-20 min. are placed onto feeder plates. After 48 h of co-incubation with *Agrobacterium*, the hypocotyl explants are transferred to B5 0/1/0 callus induction medium which contains filter sterilized carbenicillin (500 mg/l, added after autoclaving) and kanamycin sulfate (Boehringer Mannheim) at concentrations of 25 mg/l.

After 3-7 days in culture at 65 $\mu Em^{-2}S^{-1}$ to 75 $\mu Em^{-2}S^{-1}$ continuous light, callus tissue is visible on the cut surface and the hypocotyl explants are transferred to shoot induction medium, B5BZ (B5 salts and vitamins supplemented

with 3 mg/l benzylaminopurine, 1 mg/l zeatin, 1% sucrose, 0.6% Phytagar and pH adjusted to 5.8). This medium also contains carbenicillin (500 mg/l) and kanamycin sulfate (25 mg/l). Hypocotyl explants are subcultured onto fresh shoot induction medium every two weeks.

Shoots regenerate from the hypocotyl calli after one to three months. Green shoots at least 1 cm tall are excised from the calli and placed on medium containing B5 salts and vitamins, 1% sucrose, carbenicillin (300 mg/l), kanamycin sulfate (50 mg/l) and 0.6% Phytagar) and placed in a culture room with conditions as described for seed germination. After 2-4 weeks shoots which remain green are cut at the base and transferred to Magenta boxes containing root induction medium (B5 salts and vitamins, 1% sucrose, 2 mg/l indolebutyric acid, 50 mg/l kanamycin sulfate and 0.6% Phytagar). Green rooted shoots are tested for NPT II activity.

**Example 11**

In this example, a DNA-bombardment plant transformation is described. Peanut transformation is exemplified.

DNA sequences of interest may be introduced as expression cassettes, comprising at least a promoter region, a gene of interest, and a termination region, into a plant genome via particle bombardment as described in European Patent Application 332 855 and in co-pending application USSN 07/225,332, filed July 27, 1988.

Briefly, tungsten or gold particles of a size ranging from 0.5μM-3μM are coated with DNA of an expression cassette. This DNA may be in the form of an aqueous mixture or a dry DNA/particle precipitate.

Tissue used as the target for bombardment may be from cotyledonary explants, shoot meristems, immature leaflets, or anthers.

The bombardment of the tissue with the DNA-coated particles is carried out using a Biolistics™ particle gun (Dupont; Wilmington, DE). The particles are placed in the barrel at variable distances ranging from 1cm-14cm from the

barrel mouth. The tissue to be bombarded is placed beneath the stopping plate; testing is performed on the tissue at distances up to 20 cm. At the moment of discharge, the tissue is protected by a nylon net or a combination of nylon nets with mesh ranging from 10μM to 300μM.

Following bombardment, plants may be regenerated following the method of Atreya, *et al.*, (*Plant Science Letters* (1984) *34*:379-383). Briefly, embryo axis tissue or cotyledon segments are placed on MS medium (Murashige and Skoog, *Physio. Plant.* (1962) *15*:473) (MS plus 2.0 mg.l 6-benzyladenine (BA) for the cotyledon segments) and incubated in the dark for 1 week at 25 ± 2°C and are subsequently transferred to continuous cool white fluorescent light (6.8 $W/m^2$). On the 10th day of culture, the plantlets are transferred to pots containing sterile soil, are kept in the shade for 3-5 days are and finally moved to greenhouse.

The putative transgenic shoots are rooted. Integration of exogenous DNA into the plant genome may be confirmed by various methods known to those skilled in the art.

**Example 12**

This example describes methods to obtain desaturase cDNA clones from other plant species using the DNA from the *C. tinctorius* Δ-9 desaturase clone as the probe.

*Isolation of RNA for Northern Analysis*

Poly(A)+ RNA is isolated from *C. tinctorius* embryos collected at 14-17 days post-anthesis and *Simmondsia chinensis* embryos as described in Example 5.

Total RNA is isolated from days 17-18 days post-anthesis *Brassica campestris* embryos by an RNA minipreparation technique (Scherer and Knauf, *Plant Mol. Biol.* (1987) *9*:127-134). Total RNA is isolated from *R. communis* immature endosperm of about 14-21 days post-anthesis by a method described by Halling, *et al.* (*Nucl. Acids Res.* (1985) *13*:8019-8033). Total RNA is isolated

from 10 g each of young leaves from *B. campestris*, *B. napus*, and *C. tinctorius*, by extraction of each sample in 5 ml/g tissue of 4 M guanidine thiocyanate buffer as described by Colbert *et al.* (*Proc. Nat. Acac. Sci.* (1983) *80*:2248-2252). Total RNA is also isolated from immature embryos of *Cuphea hookeriana* by extraction as above in 10 ml/g tissue.

Total RNA is isolated from immature embryos of California bay (*Umbellularia californica*) by an adaptation of the method of Lagrimini *et al.* (*Proc. Nat. Acad. Sci.* (1987) *84*:7542-7546). Following homogenization in grinding buffer (2.5 ml/g tissue) as described, RNA is precipitated from the aqueous phase by addition of 1/10 volume 3 M sodium acetate and 2 volumes ethanol, followed by freezing at -80°C for 30 minutes and centrifugation at 13,000 x g for 20 minutes. The pellets are washed with 80% ethanol and centrifugation is repeated as above. The pellets are resuspended in water, two volumes of 4 M LiCl are added, and the samples are placed at -20°C overnight. Samples are centrifuged as above and the pellets washed with 80% ethanol. Ethanol precipitation is repeated as above.

Total RNA is further purified from *B. campestris*, *B. napus*, and *C. tinctorius* leaves, and from *C. tinctorius*, *B. campestris*, California bay, and jojoba, and from *R. communis* immature endosperm, by removing polysaccharides on a 0.25 g Sigma Cell 50 cellulose column. The RNA is loaded onto the column in 1 ml of loading buffer (20 mM Tris-HCl pH 7.5, 0.5M NaCl, 1mM EDTA, 0.1% SDS), eluted with loading buffer, and collected in 500 µl fractions. Ethanol is added to the samples to precipitate the RNA. The samples are centrifuged, and the pellets resuspended in sterile distilled water, pooled, and again precipitated in ethanol. The sample is centrifuged, and the resulting RNA is subjected to oligo(dT)-cellulose chromatography to enrich for poly(A)+ RNA as described by Maniatis *et al.* (*Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York (1982)). Poly(A)+ RNA is also

purified from total *Cuphea hookeriana* RNA by oligo(dT)-cellulose chromatography.

*Northern Analysis Using C. tinctorius Desaturase Clone:* 2.5 µg of poly(A)+ RNA from each of the above described poly(A)+ samples from immature embryos of jojoba, *Cuphea hookeriana*, California bay, *Brassica campestris*, and *C. tinctorius*, from immature endosperm of *R. communis*, and from leaves of *C. tinctorius*, *B. campestris*, and *B. napus* are electrophoresed on formaldehyde/agarose gels (Fourney *et al.*, *Focus* (1988) *10*:5-7) and transferred to a Hybond-C extra (Amersham, Arlington Heights, IL) filter according to manufacturer's specifications. The filter is prehybridized for four hours and hybridized overnight at 42°C in a roller bottle containing 10 ml of hybridization buffer (1 M NaCl, 1% SDS, 50% formamide, 0.1 mg/ml denatured salmon sperm DNA) in a Hybridization Incubator, model 1040-00-1 (Robbins Scientific Corporation, Sunnyvale, CA). The probe used in the hybridization is a gel-isolated *Bgl*II fragment of the Δ-9 desaturase clone that is labeled with $^{32}$P-dCTP using a BRL (Gaithersburg, MD) nick-translation kit, following manufacturer's instructions. The blot is washed three times for 20 minutes each in 2X SSC, 0.5% SDS at 55°C. The blot is exposed at -80°C, with a Dupont Cronex intensifying screen, to X-ray film for four days.

The autoradiograph shows that the *C. tinctorius* desaturase gene is expressed in both immature embryos and leaves of *C. tinctorius*, although the level of expression is considerably higher in embryos than in leaves. The autoradiograph also shows hybridization of the *C. tinctorius* desaturase clone to mRNA bands of a similar size in immature embryos from jojoba and California bay, and immature endosperm from *R. communis*. Hybridization is also detectable in RNA from *B. campestris* embryos upon longer exposure of the filter to X-ray film.

*R. communis cDNA Library Construction*: A plant seed cDNA library may be constructed from poly(A)+ RNA isolated from *R. communis* immature endosperm as described above. The plasmid cloning vector pCGN1703, and cloning method are

as described in Example 5. The *R. communis* endosperm cDNA bank contains approximately $2x10^6$ clones with an average cDNA insert size of approximately 1000 base pairs.

The *R. communis* immature endosperm cDNA bank is moved into the cloning vector lambda gt22 (Stratagene Cloning Systems) by digestion of total cDNA with *Not*I and ligation to lambda gt22 DNA digested with *Not*I. The resulting phage are packaged using a commercially available kit and titered using *E. coli* strain LE392 (Stratagene Cloning Systems, La Jolla, CA). The titer of the resulting library was approximately $1.5 \times 10^7$ pfu/ml.

*R. communis cDNA Library Screen*: The library is plated on *E. coli* strain LE392 at a density of approximately 25,000 pfu/150mm NZY plate to provide approximately 50,000 plaques for screening. Phage are lifted in duplicate on to NEN (Boston, MA.) Colony/Plaque Screen filters as described in Example 5. Following prehybridization at 42°C in 25 ml of hybridization buffer (1 M NaCl, 1% SDS, 50% formamide, 0.1 mg/ml denatured salmon sperm DNA) filters are hybridized overnight with a gel-purified 520 base pair *Bgl*II fragment of the *C. tinctorius* desaturase clone (Figure 7A) that is radiolabeled with $^{32}$P-dCTP using a BRL (Gaithersburg, MD) Nick Translation System. Filters are washed three times for 20 minutes each in 2X SSC, 0.5% SDS at 55°C in a shaking water bath. Filters are exposed to X-ray film overnight at -80°C with a Dupont Cronex intensifying screen.

Clones are detected by hybridization on duplicate filters with the *C. tinctorius* desaturase cDNA fragment and plaque purified. During plaque purification, it was observed that larger plaques were obtained when *E. coli* strain Y1090 (Young, R.A. and Davis, R.W., *Proc. Natl. Acad. Sci. USA* (1983) *80*:1194) was used as the host strain. This strain was thus used in subsequent plaque purification steps. Phage DNA is prepared from the purified clones as described by Grossberger (*NAR* (1987) *15*:6737) with the following modification. The proteinase K

treatment is replaced by the addition of 10% SDS and a 10 minute incubation at room temperature. Recovered phage DNA is digested with *Eco*RI, religated at low concentration, and transformed into *E. coli* DH5α(BRL; Gaithersburg, MD) cells to recover plasmids containing cDNA inserts in pCGN1703. Minipreparation DNA (Maniatis *et al., supra*) is prepared from the clones and DNA sequence is determined as described above. Partial nucleotide sequence of the cDNA insert of a *R. communis* desaturase clone pCGN3230 is presented in Figure 3A and SEQ ID NO: 14. The complete nucleotide sequence of this clone is presented in Fig. 3B and as SEQ ID NO: 15.

*Northern Analysis Using R. communis Desaturase Clone*: Total RNA for Northern analysis is isolated from tobacco leaves by the method of Ursin *et al.*(*Plant Cell* (1989) *1*:727-736), petunia and tomato leaves by the method of Ecker and Davis (*Proc.Nat.Acad.Sci.* (1987) *84*:5202-5206), and corn leaves by the method of Turpen and Griffith (*Biotechniques* (1986) *4*:11-15). Total RNA samples from tobacco, corn, and tomato leaves are enriched for poly(A)+ RNA by oligo(dT)-cellulose chromatography as described by Maniatis *et al.* (*supra*).

Poly(A)+ RNA samples from tomato leaves (4 μg) and corn and tobacco leaves (1 μg each), and total RNA from petunia leaves (25 μg) are electrophoresed on a formaldehyde/agarose gel as described by Shewmaker *et al.* (*Virology* (1985) *140*:281-288). Also electrophoresed on this gel are poly(A)+ RNA samples isolated from *B. campestris* day 17-19 embryos and *B. campestris* leaves (2 μg each), immature embryos from *C. tinctorius*, bay, and jojoba (1 μg each), and *R. communis* endosperm (1 μg). The isolation of these poly(A)+ RNA samples is described above for the Northern analysis using *C. tinctorius* desaturase cDNA as probe. The RNA is transferred to a nitrocellulose filter as described by Shewmaker *et al.* (*supra*) and prehybridized and hybridized at 42°C in 50% formamide, 10X Denhardt's solution (described in Maniatis *et al.* (supra)), 5X SSC, 0.1% SDS, 5 mM EDTA, 100 ug/ml denatured salmon

sperm DNA, and 10% dextran sulfate (in hybridization buffer only). The probe for hybridization is the $^{32}P$-labeled (BRL Nick Translation Kit) 1.7 kb SalI insert of pCGN3230 that has been gel-purified from minipreparation DNA. The filter is washed following hybridization for 30 minutes in 2X SSC, 0.1% SDS at 42°C and at 50°C twice for 15 minutes each. The filter is exposed to X-ray film overnight at -80°C with a Dupont Cronex intensifying screen.

The autoradiograph shows hybridization of the *R. communis* desaturase clone to mRNA bands of a similar size in immature embryos from *B. campestris*, California bay, and *C. tinctorius*, and also in corn leaves and *R. communis* endosperm.

*B. campestris Embryo cDNA Library Construction:* Total RNA is isolated from 5 g of *B. campestris* cv. R500 embryos obtained from seeds harvested at days 17-19 post-anthesis. RNA is extracted in 25 mls of 4 M guanidine thiocyanate buffer as described by Colbert *et al.*(*PNAS* (1983) *80*:2248-2252). Polysaccharides are removed from the RNA sample by resuspending the pellet in 6 ml of 1X TE (10 mM Tris/1 mM EDTA pH 8), adding potassium acetate to a concentration of 0.05M, and adding one half volume of ethanol. The sample is placed on ice for 60 minutes and centrifuged for 10 minutes at 3000 x g. RNA is precipitated from the supernatant by adding sodium acetate to a concentration of 0.3 M followed by the addition of two volumes of ethanol. RNA is recovered from the sample by centrifugation at 12,000 x g for 10 minutes and yield calculated by UV spectrophotometry. Two mg of the total RNA is further purified by removing polysaccharides on a 0.25 g Sigma Cell 50 cellulose column, as described above, and is also enriched for poly(A)+ RNA by oligo(dT)-cellulose chromatography as described above.

A *B. campestris* day 17-19 post anthesis embryo cDNA library is constructed in plasmid vector pCGN1703 as described in Example 5, using 5 ug of the above described poly(A)+ RNA. The library, which consists of approximately 1.5 x $10^5$ transformants, is amplified by plating and scraping colonies, and is stored as frozen *E. coli* cells in

10% DMSO at -80° C. DNA is isolated from a portion of the amplified library by scaling up the alkaline lysis technique of Birnboim and Doly (*Nucleic Acids Res.* (1979) 7:1513), and purified by CsCl centrifugation. Library DNA is digested with *Eco*RI and is cloned into *Eco*RI-digested bacteriophage lambda gt10 (Stratagene; La Jolla, CA) DNA. The DNA is packaged using Gigapack II Gold *in vitro* packaging extracts (Stratagene; La Jolla, CA) according to manufacturer's specifications. The titer of the phage stock, determined by dilution plating of phage in *E. coli* C600 hfl- cells (Huynh, *et al., DNA Cloning. Volume 1.* Eds. Gover, D.M. (1985) IRL Press Limited: Oxford, England, pp. 56,110), is 6 x $10^6$ pfu per ml.

*B. campestris cDNA Library Screen*: The library is plated on *E. coli* strain C600 hfl- at a density of approximately 30,000 pfu/150mm NZY plate to provide approximately 120,000 plaques for screening. Phage are lifted in duplicate on to NEN (Boston, MA.) Colony/Plaque Screen filters as described in Example 5. Filters are prehybridized and hybridized with the $^{32}$P-labeled fragment of pCGN3230 as described above for the Northern hybridization. Filters are washed for 30 minutes in 2X SSC, 0.1% SDS at 50°C and at 55°C twice for 15 minutes each. Filters are exposed to X-ray film overnight at -80°C with a Dupont Cronex intensifying screen.

Clones are detected by hybridization on duplicate filters to the *R. communis* desaturase cDNA fragment and plaque purified. During plaque purification, the probe used was a gel-purified 1.4 kb SstI fragment of pCGN3230 which lacks the poly(A)+ tail. As described above, phage DNA is isolated from purified lambda clones, digested with EcoRI, ligated, and transformed to *E. coli* DH5α cells. Minipreparation DNA is prepared and partial DNA sequence determined as described above. Partial DNA sequences of two clones, pCGN3235 and pCGN3236, are presented in Figure 4A (SEQ ID NO: 17) and 4B (SEQ ID NO: 18), respectively. Initial DNA sequence analysis of the 3' regions of these clones indicates that pCGN3236 and pCGN3235 are cDNA

clones from the same gene. pCGN3236 is a shorter clone than pCGN3235, which appears to contain the entire coding region of the *B. campestris* desaturase gene. The complete nucleotide sequence of pCGN3235 is presented in Figure 4C and SEQ ID NO: 19.

*Desaturase Gene Analysis*: Southern and Northern analyses of Brassica species are conducted to determine the number of genes which encode the Brassica desaturase clone, pCGN3235 in *B. campestris*, *B. oleracea*, and *B. napus*, and the timing of expression of the gene in *B. campestris* developing seeds. DNA is isolated from leaves of each of the above-named *Brassica* species by the method of Bernatzky and Tanksley (*Theor. Appl. Genet.* (1986) 72:314-321). DNA from each of the species is digested with restriction endonucleases EcoRI and XbaI (10 ug/digest), electrophoresed in a 0.7% agarose gel, and transferred to a nitrocellulose filter (Maniatis *et al.*, *supra*). The filter is prehybridized and hybridized at 42°C (as described above for Northern analysis using *R. communis* desaturase clone) with a $^{32}$P-labeled (nick translation) gel-isolated HindIII/PvuII fragment of pCGN3235 (Fig. 7C). The filter is washed following overnight hybridization, for 30 minutes at 55°C in 1X SSC, 0.1% SDS, followed by two 15 minute washed at 55°C in 0.1X SSC, 0.1% SDS.

The autoradiograph indicates that the *Brassica* desaturase is encoded by a small gene family consisting of about two genes in *B. campestris* and *B. oleracea*, and about four genes in *B. napus*.

The timing of expression of the desaturase gene during seed development is determined by Northern analysis. RNA is isolated from immature seeds of *B. campestris* cv. R500 collected at 11, 13, 15, 17, 19, 21, 25, 30, 35, and 40 days post-anthesis. Total RNA is isolated as described by Scherer and Knauf (*Plant Mol. Biol.* (1987) *9*:127-134). Twenty five micrograms of RNA from each time point are electrophoresed through a formaldehyde-containing 1.5% agarose gel as described by Shewmaker, *et al.* (*supra*) and blotted to nitrocellulose (Thomas, *Proc. Nat. Acad. Sci.*

(1980) 77:5201-5205). The blot is pre-hybridized and hybridized at 42°C with the $^{32}$P-labeled HindIII/PvuII fragment of pCGN3235 as described above. The filter is washed following overnight hybridization, for 30 minutes at 55°C in 1X SSC, 0.1% SDS, followed by two 15 minute washed at 55°C in 0.1X SSC, 0.1% SDS.

The autoradiograph indicates that the desaturase gene is expressed in *B. campestris* developing seeds beginning at about day 19 and through about day 30, with maximal expression at day 25. By a similar Northern analysis, the level of desaturase mRNA in developing *Brassica napus* seeds (day 21) was estimated to be approximately 1% of the total mRNA.

*Isolation of Other Desaturase Gene Sequences:* cDNA libraries may be constructed as described above and genomic libraries can be constructed from DNA from various sources using commercially available vectors and published DNA isolation, fractionation, and cloning procedures. For example, a *B. campestris* genomic library can be constructed using DNA isolated according to Scofield and Crouch (J.Biol.Chem. (1987) 262:12202-12208) that is digested with BamHI and fractionated on sucrose gradients (Maniatis *et al., supra*), and cloned into the lambda phage vector LambdaGem-11 (Promega; Madison,WI) using cloning procedures of Maniatis *et al.* (*supra*).

cDNA and genomic libraries can be screened for desaturase cDNA and genomic clones, respectively, using published hybridization techniques. Screening techniques are described above for screening libraries with DNA fragments. Libraries may also be screened with synthetic oligonucleotides, for example using methods described by Berent *et al.* (*BioTechniques* (1985) 3:208-220). Probes for the library screening can be prepared by PCR, or from the sequences of the desaturase clones provided herein. Oligonucleotides prepared from the desaturase sequences may be used, as well as longer DNA fragments, up to the entire desaturase clone.

For example, jojoba polyadenylated RNA is used to construct a cDNA library in the cloning vector λZAPII/*Eco*RI (Stratagene, San Diego, CA). RNA is isolated from jojoba embryos collected at 80-90 days post-anthesis by isolating polyribosomes using a method initially described by Jackson and Larkins (*Plant Physiol.* (1976) *57*:5-10) and modified by Goldberg *et al.* (*Developmental Biol.* (1981) *83*:201-217). Polysaccharide contaminants in the polyribosomal RNA preparation are removed by running the RNA over a cellulose column (Sigma-cell 50) in high salt buffer (0.5M NaCl, 20mM Tris pH 7.5, 1mM EDTA, 0.1% SDS). The contaminant binds to the column and the RNA is collected in the eluant. The eluant fractions are pooled and the RNA is ethanol precipitated. The precipitated total RNA is then resuspended in a smaller volume and applied to an oligo d(T) cellulose column to isolate the polyadenylated RNA.

The library is constructed using protocols, DNA and bacterial strains as supplied by the manufacturer. Clones are packaged using Gigapack Gold packaging extracts (Stratagene), also according to manufacturer's recommendations. The cDNA library constructed in this manner contins approximately 1 x $10^6$ clones with an average cDNA insert size of approximately 400 base pairs.

The jojoba library is plated on *E. coli* XL1-Blue (Stratagene) at a density of approximately 5000pfu/150mm plate to provide approximately 60,000 plaques for screening. Phage are lifted onto duplicate nylon membrane filters as described previously. Filters are prehybridized at 42°C in a hybridization buffer containing 40% formamide, 10X Denhardt's solution, 5X SSC, 0.1% SDS, 50mM EDTA, and 100µg/ml denatured salmon sperm DNA. Hybridization is at 42°C in the same buffer with added nick translated (BRL Nick Translation System) 520 bp *Bgl*II fragment of the *C. tinctorius* desaturase clone described previously. Filters are washed at 50°C in 2X SSC and exposed to X-ray film overnight.

Desaturase clones are detected by hybridization on duplicate filters with the *C. tinctorius* cDNA fragment and

plaque-purified. Positive clones are recovered as plasmids in *E. coli* following manufacturer's directions and materials for *in vivo* excision. Partial, preliminary DNA sequence of a clone, 3-1, is determined and the corresponding amino acid sequence is translated in three frames. In this manner, homology to the *C. tinctorius* desaturase cDNA clone is detected in one reading frame. The preliminary DNA sequence of this jojoba desaturase cDNA fragment is shown in Figure 5 (SEQ ID NO: 43). Also shown is the corresponding translated amino acid sequence in the reading frame having *C. tinctorius* desaturase homology. The jojoba cDNA fragment is approximately 75% homologous at the DNA level and approximately 79% homologous at the amino acid level compared to sequence of the *C. tinctorius* desaturase in this region.

**Example 13**

Antisense constructs are described which allow for transcription of a reverse copy of the *B. campestris* desaturase cDNA clone in the 5' to 3' orientation of transcription.

*Preferential Expression of Antisense Constructs in Embryos*

In order to reduce the transcription of a desaturase gene in embryos of *B. napus* or *B. campestris*, constructs may be prepared which allow for production of antisense copies of the desaturase cDNA preferentially in the embryos. Promoter sequences which are desirable to obtain this pattern of expression include, but are not limited to, the ACP, Bce4, and napin 1-2 expression cassettes described in Examples 7, 8, and 9, respectively. It also may be desirable to control the expression of reverse copies of the desaturase cDNA under two different promoters in the same transformed plant to provide for a broader timing of expression of the antisense desaturase DNA. For example, expression from the ACP promoter may begin and end earlier than expression from the napin promoter. Thus, expressing the reverse desaturase from both promoters may result in

the production of the antisense strand of DNA over a longer period of embryo development.

An example of expression of an antisense desaturase gene preferentially in the embryos is provided below. Similar constructs containing the same or a different fragment of the desaturase gene and any of the promoters described above, as well as other promoter regions which may be useful, may also be prepared using gene cloning, insertion, mutation and repair techniques well known to those of ordinary skill in the art.

A. Antisense Desaturase Expression from the ACP Promoter

Construction of pCGN3239 is as follows:

pCGN3235 (Example 12) is digested with *Pvu*II and *Hin*dIII and the *Hin*dIII sticky ends are filled in with Klenow in the presence of 200 μM dNTPs. The 1.2 kb *Pvu*II/*Hin*dIII fragment containing the desaturase coding sequence is gel purified and ligated in the antisense orientation into *Eco*RV-digested pCGN1977 (ACP expression cassette; described in Example 7) to create pCGN3238. The 4.2 kb *Xba*I/*Asp*718 fragment of pCGN3238 containing the antisense desaturase in the ACP cassette is transferred into *Xba*I/*Asp*718-digested pCGN1557 (binary transformation vector; described in Example 7) to create pCGN3239.

B. Antisense Desaturase Expression From The Napin Promoter

Construction of pCGN3240 is as follows: pCGN3235 is digested with *Pvu*II and *Hin*dIII, the sticky ends are blunted, and the resulting fragment is inserted in an anti-sense orientation into pCGN3223 which has been digested with *Sal*I and blunted with Klenow enzyme. The resulting plasmid, pCGN3240 will express an anti-sense desaturase RNA from the napin promoter cassette.

C. Antisense Desaturase Expression From a Dual Promoter Cassette

Construction of pCGN3242 is as follows: An *Asp*718 fragment of pCGN3240 containing the napin 5' and 3' regions surrounding the desaturase sequences is inserted into the

*Asp*718 site of pCGN3239 (a binary vector containing an ACP promoter, antisense desaturase construct) to create pCGN3242.

*Constitutive Transcription*

A. Binary Vector Construction

1. Construction of pCGP291.

The *Kpn*I, *Bam*HI, and *Xba*I sites of binary vector pCGN1559 (McBride and Summerfelt, *Pl.Mol.Biol.* (1990) *14*: 269-276) are removed by *Asp*718/*Xba*I digestion followed by blunting the ends and recircularization to produce pCGP67. The 1.84 kb *Pst*I/*Hin*dIII fragment of pCGN986 containing the 35S promoter-tml3' cassette is inserted into *Pst*I/*Hin*dIII digested pCGP67 to produce pCGP291.

2. Construction of pCGN986.

The 35S promoter-tml3' expression cassette, pCGN986, contains a cauliflower mosaic virus 35S (CaMV35) promoter and a T-DNA tml 3'-region with multiple restriction sites between them. pCGN986 is derived from another cassette, pCGN206, containing a CaMV35S promoter and a different 3' region, the CaMV region VI 3'-end. The CaMV 35S promoter is cloned as an *Alu*I fragment (bp 7144-7734) (Gardner *et. al.*, *Nucl.Acids Res.* (1981) *9*:2871-2888) into the *Hin*cII site of M13mp7 (Messing, *et. al.*, *Nucl.Acids Res.* (1981) *9*:309-321) to create C614. An *Eco*RI digest of C614 produced the *Eco*RI fragment from C614 containing the 35S promoter which is cloned into the *Eco*RI site of pUC8 (Vieira and Messing, *Gene* (1982) *19*:259) to produce pCGN147.

pCGN148a containing a promoter region, selectable marker (KAN with 2 ATG's) and 3' region, is prepared by digesting pCGN528 with *Bgl*II and inserting the *Bam*HI-*Bgl*II promoter fragment from pCGN147. This fragment is cloned into the *Bgl*II site of pCGN528 so that the *Bgl*II site is proximal to the kanamycin gene of pCGN528.

The shuttle vector used for this construct, pCGN528, is made as follows: pCGN525 is made by digesting a plasmid containing Tn5 which harbors a kanamycin gene (Jorgenson

*et. al., Mol. Gen. Genet.* (1979) *177*:65) with *Hin*dIII-*Bam*HI and inserting the *Hin*dIII-*Bam*HI fragment containing the kanamycin gene into the *Hin*dIII-*Bam*HI sites in the tetracycline gene of pACYC184 (Chang and Cohen, *J. Bacteriol.* (1978) *134*:1141-1156). pCGN526 was made by inserting the *Bam*HI fragment 19 of pTiA6 (Thomashow *et. al., Cell* (1980) *19*:729-739), modified with *Xho*I linkers inserted into the *Sma*I site, into the *Bam*HI site of pCGN525. pCGN528 is obtained by deleting the small *Xho*I fragment from pCGN526 by digesting with *Xho*I and religating.

pCGN149a is made by cloning the *Bam*HI-kanamycin gene fragment from pMB9KanXXI into the *Bam*HI site of pCGN148a. pMB9KanXXI is a pUC4K variant (Vieira and Messing, *Gene* (1982) *19*:259-268) which has the *Xho*I site missing, but contains a functional kanamycin gene from Tn903 to allow for efficient selection in *Agrobacterium*.

pCGN149a is digested with *Hin*dIII and *Bam*HI and ligated to pUC8 digested with *Hin*dIII and *Bam*HI to produce pCGN169. This removes the Tn903 kanamycin marker. pCGN565 and pCGN169 are both digested with *Hin*dIII and *Pst*I and ligated to form pCGN203, a plasmid containing the CaMV 35S promoter and part of the 5'-end of the Tn5 kanamycin gene (up to the *Pst*I site, Jorgenson *et. al.*, (1979), *supra*). A 3'-regulatory region is added to pCGN203 from pCGN204, an *Eco*RI fragment of CaMV (bp 408-6105) containing the region VI 3' cloned into pUC18 (Yanisch-Perron, *et al., Gene* (1985) *33*:103-119) by digestion with *Hin*dIII and *Pst*I and ligation. The resulting cassette, pCGN206, is the basis for the construction of pCGN986.

The pTiA6 T-DNA tml 3'-sequences are subcloned from the *Bam*19 T-DNA fragment (Thomashow *et al.*, (1980) *supra*) as a *Bam*HI-*Eco*RI fragment (nucleotides 9062 to 12,823, numbering as in Barker *et al., Plant Mol. Biol.* (1982) *2*:335-350) and combined with the pACYC184 (Chang and Cohen (1978), *supra*) origin of replication as an *Eco*RI-*Hin*dIII fragment and a gentamycin resistance marker (from plasmid

pLB41), obtained from D. Figurski) as a *Bam*HI-*Hin*dIII fragment to produce pCGN417.

The unique *Sma*I site of pCGN417 (nucleotide 11,207 of the *Bam*19 fragment) is changed to a *Sac*I site using linkers and the *Bam*HI-*Sac*I fragment is subcloned into pCGN565 to give pCGN971. The *Bam*HI site of pCGN971 is changed to an *Eco*RI site using linkers. The resulting *Eco*RI-*Sac*I fragment containing the tml 3' regulatory sequences is joined to pCGN206 by digestion with *Eco*RI and *Sac*I to give pCGN975. The small part of the Tn5 kanamycin resistance gene is deleted from the 3'-end of the CaMV 35S promoter by digestion with *Sal*I and *Bgl*II, blunting the ends and ligation with *Sal*I linkers. The final expression cassette pCGN986 contains the CaMV 35S promoter followed by two *Sal*I sites, an *Xba*I site, *Bam*HI, *Sma*I, *Kpn*I and the tml 3' region (nucleotides 11207-9023 of the T-DNA).

B. Insertion of Desaturase Sequence

The 1.6 kb *Xba*I fragment from pCGN3235 containing the desaturase cDNA is inserted in the antisense orientation into the *Xba*I site of pCGP291 to produce pCGN3234.

*Plant Transformation*

The binary vectors containing the expression cassette and the desaturase gene are transformed into *Agrobacterium tumefaciens* strain EHA101 (Hood, *et al., J. Bacteriol.* (1986) *168*:1291-1301) as per the method of Holsters, *et al., Mol. Gen. Genet.* (1978) *163*:181-187. Transformed *B. napus* and/or *Brassica campestris* plants are obtained as described in Example 10.

*Analysis of Transgenic Plants*

A. Analysis of pCGN3242 Transformed *Brassica campestris* cv. Tobin Plants

Due to the self-incompatibility of *Brassica campestris* cv. Tobin, individual transgenic plants are pollinated using non-transformed Tobin pollen. Because of this, the T2 seeds of a transgenic plant containing the antisense desaturase at one locus would be expected to segregate in a

1:1 ratio of transformed to non-transformed seed. The oil composition of ten individual seeds collected at 26 days post-anthesis from several pCGN3242-transformed plants and one non-transformed control was analyzed by gas chromatography according to the method of Browse, *et al.*, *Anal. Biochem.* (1986) *152*:141-145. One transformant, 3242-T-1, exhibits an oil composition that differed distinctly from controls on preliminary analysis. The control Tobin seeds contained an average of 1.8% 18:0 (range 1.5% - 2.0%) and 52.9% 18:1 (range 48.2% - 57.1%). T2 seeds of 3242-T-1 segregated into two distinct classes. Five seeds contained levels of 18:0 ranging from 1.3% to 1.9% and levels of 18:1 ranging from 42.2% to 58.3%. The other five seeds contained from 22.9% to 26.3% 18:0 and from 19.9% to 26.1% 18:1.

B. Analysis of pCGN3234 Transformed Plants

Some abnormalities have been observed in some transgenic *Brassica napus* cv. Delta and Bingo and *Brassica campestris* cv. Tobin plants containing pCGN3234. These effects could be due to the constitutive expression of antisense desaturase RNA from the 35S promoter or could be due to the transformation/tissue culture regime the plants have been subjected to.

The above results demonstrate the ability to obtain plant Δ-9 desaturases, isolate DNA sequences which encode desaturase activity and manipulate them. In this way, the production of transcription cassettes, including expression cassettes can be produced which allow for production, including specially differentiated cell production of the desired product. A purified *C. tinctorius* desaturase is provided and used to obtain nucleic acid sequences of *C. tinctorius* desaturase. Other plant desaturase sequences are provided such as *R. cummunis*, *B. campestris*, and *S. chinensis*. These sequences as well as desaturase sequences obtained from them may be used to obtain additional desaturease, and so on. And, as described in the application modification of oil composition may be achieved.

All publications and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claim.

What is claimed is:

1. A recombinant DNA construct comprising a sequence encoding at least a portion of a plant desaturase, said desaturase when mature having activity toward an unsaturated fatty acid substrate.

2. The construct of Claim 1 encoding a biologically active plant desaturase.

3. The construct of Claim 1 wherein said sequence encodes a precursor desaturase.

4. The construct of Claim 1 wherein said sequence encodes a mature desaturase.

5. The construct of Claim 1 wherein said sequence encodes a transit peptide.

6. The construct of Claim 1 comprising a cDNA sequence.

7. The construct of Claim 1 wherein said sequence is joined to a second nucleic acid sequence which is not naturally joined to said first sequence.

8. The construct of Claim 1 comprising, in the 5' to 3' direction of transcription, a transcriptional regulatory region functional in a host cell and said sequence.

9. The construct of Claim 8 further comprising, a translational regulatory region immediately 5' to said sequence and a transcriptional/translational termination regulatory region 3' to said sequence, wherein said regulatory regions are functional in said host cell.

10. The construct of Claim 8 wherein said sequence is a sense sequence.

11. The construct of Claim 8 wherein said sequence is an anti-sense sequence.

12. The construct of Claim 8 wherein said host cell is a plant cell.

13. The construct of Claim 12 wherein said transcriptional initiation region is obtained from a gene preferentially expressed in plant seed tissue during lipid accumulation.

14. The construct of Claim 13 wherein said transcriptional initiation region is selected from the

regulatory region 5' upstream to a structural gene of the group consisting of any one of Bce4, seed ACP Bcg 4-4 and napin 1-2.

15. The construct of Claim 9 wherein said transcriptional termination region is a plant desaturase termination region.

16. The construct of Claim 1 wherein said plant desaturase is a Δ-9 desaturase.

17. The construct of Claim 1 wherein said sequence is obtainable from any one of *C. tinctorius, R. communinis* and *B. campestris*.

18. A method of modifying fatty acid composition in a plant host cell from a given percentage of fatty acid saturation to a different percentage of fatty acid saturation comprising

growing a host plant cell having integrated into its genome a recombinant DNA sequence encoding a plant desaturase under the control of regulatory elements functional in said plant cell during lipid accumulation under conditions which will promote the activity of said regulatory elements.

19. The method of Claim 18 wherein the overexpression of plant desaturase is obtained.

20. The method of Claim 18 wherein the decrease of endogenous plant desaturase is obtained.

21. The method of Claim 18 wherein said regulatory elements function preferentially in plant seed.

22. The method of Claim 20 wherein the percentage of long chain unsaturated fatty acids is increased.

23. A plant cell having a modified level of saturated fatty acids produced according to the method of any one of Claims 18-22.

24. The plant cell of Claim 23 wherein said cell is a *Brassica* plant cell.

25. The plant cell of Claim 23 wherein said cell is *in vivo*.

26. The plant cell of Claim 23 wherein said cell is an oilseed embryo plant cell.

27. A plant seed having a modified level of saturated fatty acids as compared to a seed of said plant having a native level of saturated fatty acids produced according to a method comprising

growing a plant, having integrated into the genome of embryo cells a recombinant DNA sequence encoding a plant desaturase under the control of regulatory elements functional in seed during lipid accumulation, to produce seed under conditions which will promote the activity of said regulatory elements, and

harvesting said seed.

28. The seed of Claim 27 wherein said plant is *Brassica napus*.

29. The seed of claim 27 wherein said seed is an oilseed.

30. The seed of Claim 27 wherein said plant desaturase is a Δ-9 desaturase.

31. A plant seed oil of a plant having an endogenous level of saturated fatty acids comprising a plant seed oil having a modified level of saturated fatty acids.

32. The oil of Claim 31 comprising a *Brassica napus* oil.

33. A plant seed oil separated from an seed produced according to any one of Claims 27-30.

34. A host cell comprising a plant desaturase encoding sequence of any one of Claims 1-17.

35. The cell of Claim 34 wherein said cell is a plant cell.

36. The cell of Claim 35 wherein said plant cell is *in vivo*.

37. The cell of Claim 35 wherein said plant cell is a *Brassica* plant cell.

38. A transgenic host cell comprising an expressed plant desaturase.

39. The cell of Claim 38 wherein said host cell is a plant cell.

40. The cell of Claim 38 wherein said plant desaturase is a Δ-9 desaturase.

41. A method of producing a plant desaturase in a host cell or progeny thereof comprising

growing a host cell or progeny thereof comprising a construct of any one of Claims 1-10 and 12-17 under conditions which will permit the production of said plant desaturase.

42. The method of Claim 41 wherein said host cell is a plant cell and said construct is integrated into the genome of said plant cell.

43. The method of Claim 42 wherein said plant cell is *in vivo*.

44. A host cell comprising a plant desaturase produced according to Claim 41.

45. The cell of Claim 45 wherein said host cell is a plant host cell and said construct is integrated into the genome of said plant cell.

F1: ASTLGSSTPKVDNAKKPFQPPREVHVQVTH$^{S}_{X}$MPPQKIEIFKSIEG$^{W}_{R}$AEQNILV$^{H}_{F}$LKPVEKCWQ

F2: DFLPDPA$^{S}_{T}$EGFDEQVKELRARAKEIPDDYFVVLVGDMITEEALPTYQTMLNTLDGV

F3: DETGASLTPWAVWT

F4: DLLHTYLYLSGRV

F5: DMRQIQKTIQYLI

F6: TENSPYLGFIYTSFQER

F7: DV$^{K}_{F}$LAQI$^{C}_{Q}$GTIASDEKRHETAYTKIVEKLFEIDPDGTVLAFADMMRKKI$^{S}_{T}$MPAHLMY

F8: DNLF

F9: dvF1AV$^{A}_{I}$QRL$^{G}_{I}$VYTAK

F10: DYADILEFLVGRWK

F11: VADLTGLSGEGRKA$^{Q}_{G}$DYVCGLPPRIRRLEERAQGRAKEGPVVPFSWIFDRQVKL

FIGURE 1

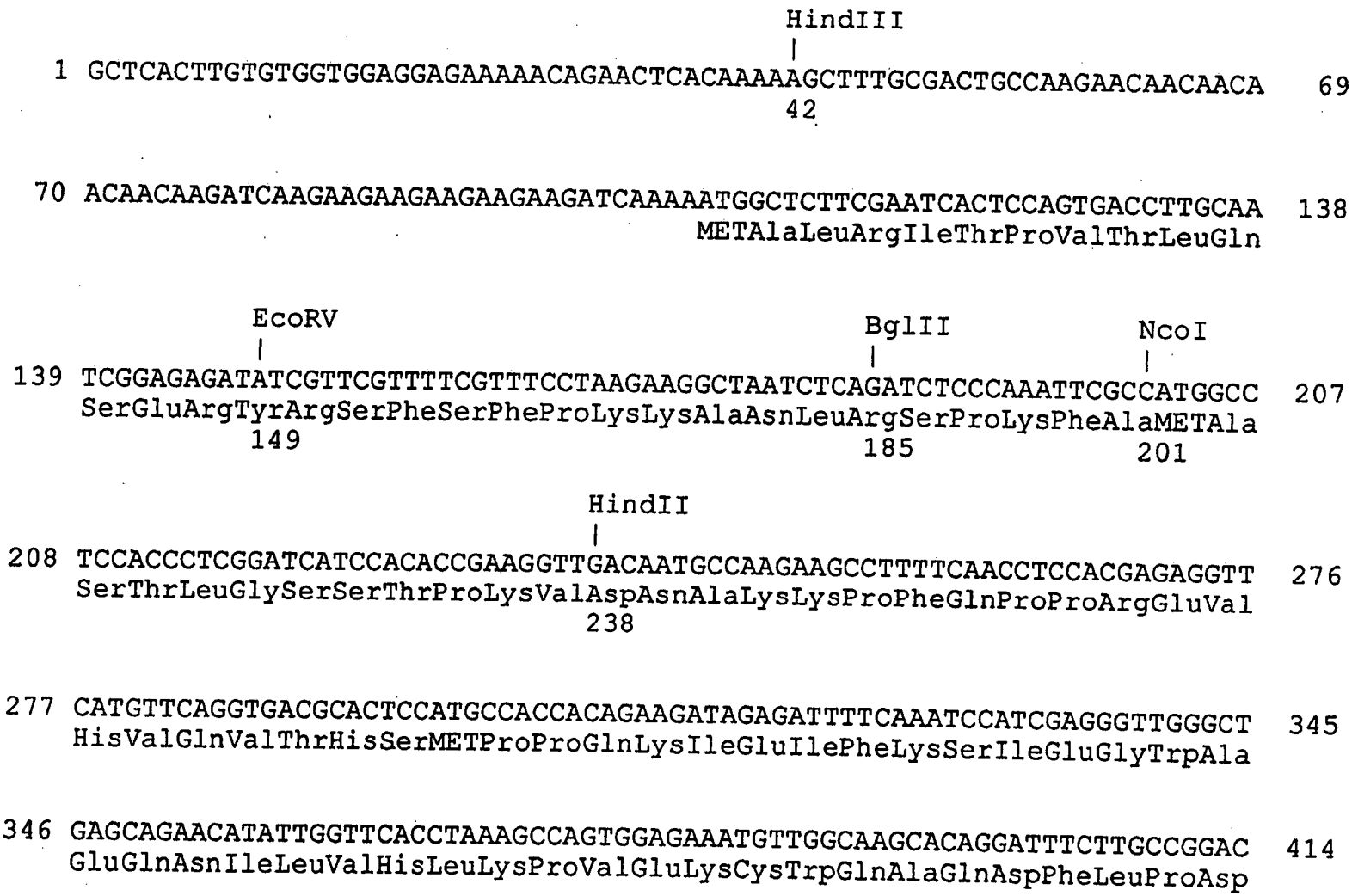



FIGURE 2
Page 1 of 4

```
415 CCTGCATCTGAAGGATTTGATGAACAAGTCAAGGAACTAAGGGCAAGAGCAAAGGAGATTCCTGATGAT  483
    ProAlaSerGluGlyPheAspGluGlnValLysGluLeuArgAlaArgAlaLysGluIleProAspAsp

484 TACTTTGTTGTTTTGGTTGGAGATATGATTACAGAGGAAGCCCTACCTACTTACCAAACAATGCTTAAT  552
    TyrPheValValLeuValGlyAspMETIleThrGluGluAlaLeuProThrTyrGlnThrMETLeuAsn

553 ACCCTAGATGGTGTACGTGATGAGACTGGGGCTAGCCTTACGCCTTGGGCTGTCTGGACTAGGGCTTGG  621
    ThrLeuAspGlyValArgAspGluThrGlyAlaSerLeuThrProTrpAlaValTrpThrArgAlaTrp

        PvuII                                                        AccI
        |                                                            |
622 ACAGCTGAAGAGAACAGGCATGGCGATCTTCTCCACACCTATCTCTACCTTTCTGGGCGGGTAGACATG  690
    ThrAlaGluGluAsnArgHisGlyAspLeuLeuHisThrTyrLeuTyrLeuSerGlyArgValAspMET
        626                                                          684

                                                BamHI
                                                |
691 AGGCAGATACAGAAGACAATTCAGTATCTCATTGGGTCAGGAATGGATCCTCGTACCGAAAACAGCCCC  759
    ArgGlnIleGlnLysThrIleGlnTyrLeuIleGlySerGlyMETAspProArgThrGluAsnSerPro
                                                736

760 TACCTTGGGTTCATCTACACATCGTTTCAAGAGCGTGCCACATTTGTTTCTCACGGAAACACCGCCAGG  828
    TyrLeuGlyPheIleTyrThrSerPheGlnGluArgAlaThrPheValSerHisGlyAsnThrAlaArg
```

FIGURE 2
Page 2 of 4

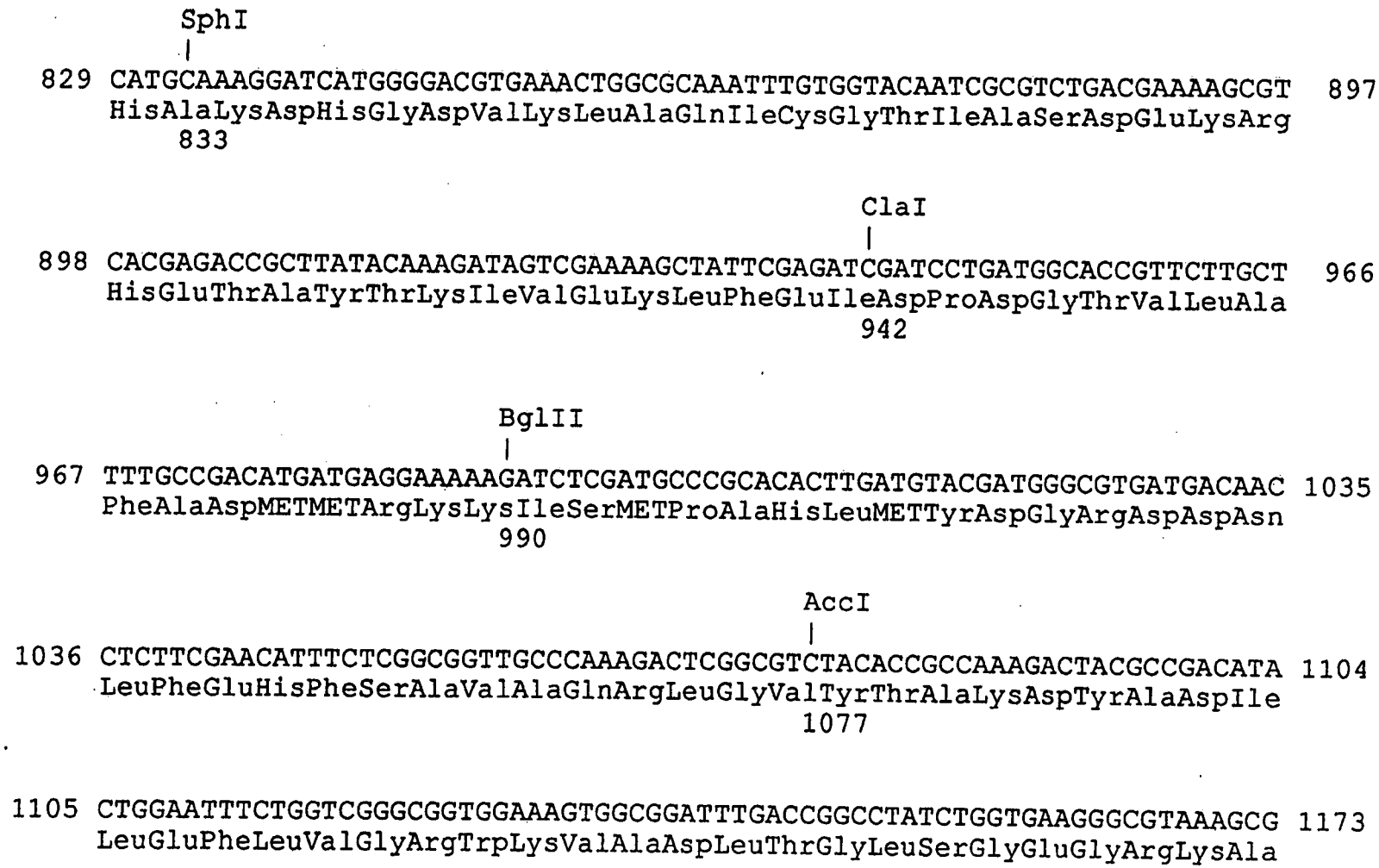



FIGURE 2
Page 3 of 4

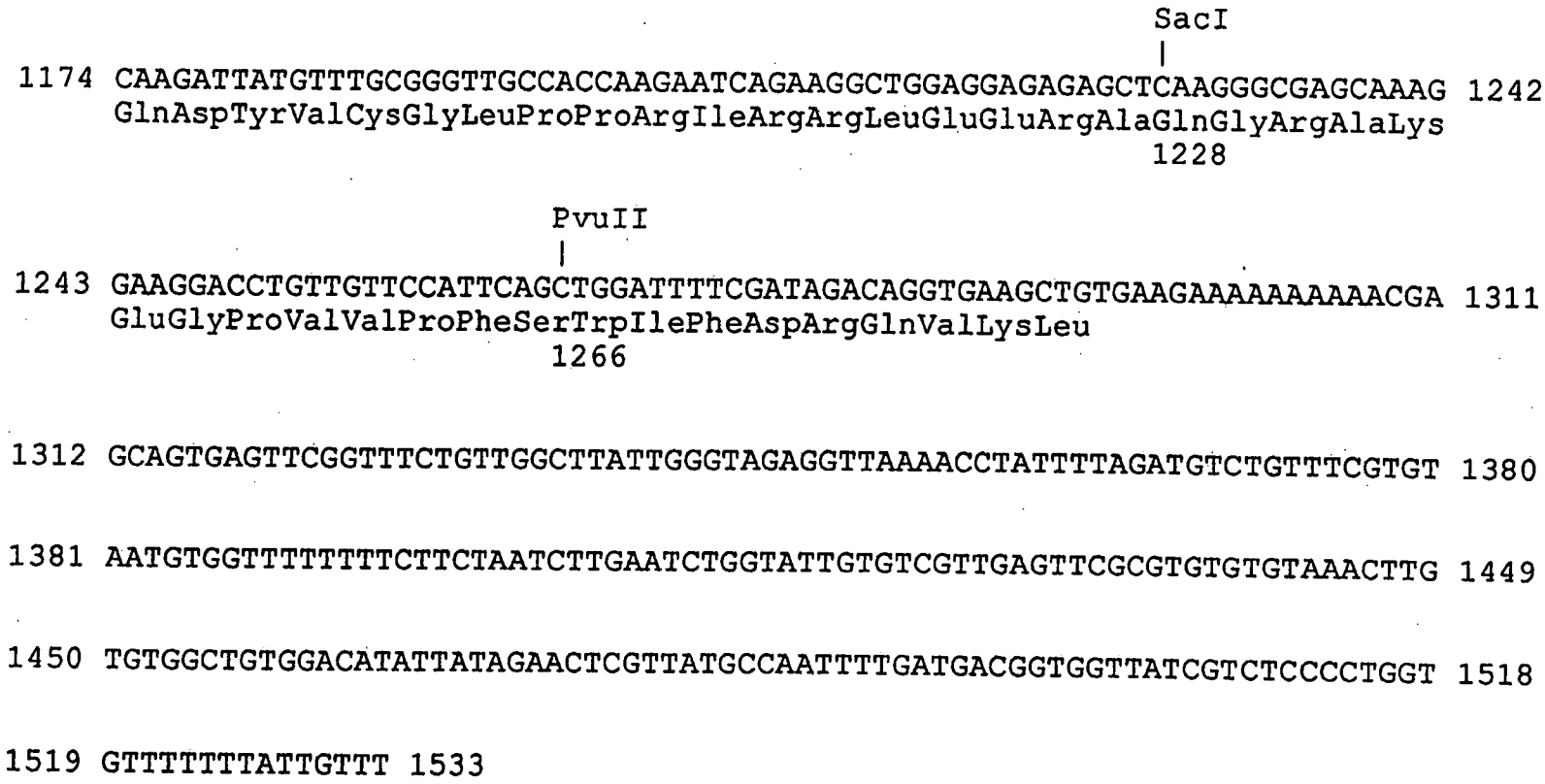



FIGURE 2
Page 4 of 4

```
1 AAAAGAAAAAGGTAAGAAAAAAAACAATGGCTCTCAAGCTCAATCCTTTCCTTTCTCAAACCCAAAAGT  69
                                METAlaLeuLysLeuAsnProPheLeuSerGlnThrGlnLysL

                                        BglII
                                        |
70 TACCTTCTTTCGCTCTTCCACCAATGGCCAGTACCAGATCTCCTAAGTTCTACATGGCCTCTACCCTCA 138
   euProSerPheAlaLeuProProMETAlaSerThrArgSerProLysPheTyrMETAlaSerThrLeuL

139 AGTCTGGTTCTAAGGAAGTTGAGAATCTCAAGAAGCCTTTCATGCCTCCTCGGGAGGTACATGTTCAGG 207
    ysSerGlySerLysGluValGluAsnLeuLysLysProPheMETProProArgGluValHisValGlnV

208 TTACCCATTCTATTGCCA 225
    alThrHisSerIleAla
```

FIGURE 3A

```
AAAAGAAAAA GGTAAGAAAA AAAACA ATG GCT CTC AAG CTC AAT CCT TTC CTT TCT          56
                             MET Ala Leu Lys Leu Asn Pro Phe Leu Ser

CAA ACC CAA AAG TTA CCT TCT TTC GCT CTT CCA CCA ATG GCC AGT ACC AGA TCT     110
Gln Thr Gln Lys Leu Pro Ser Phe Ala Leu Pro Pro MET Ala Ser Thr Arg Ser

CCT AAG TTC TAC ATG GCC TCT ACC CTC AAG TCT GGT TCT AAG GAA GTT GAG AAT     164
Pro Lys Phe Tyr MET Ala Ser Thr Leu Lys Ser Gly Ser Lys Glu Val Glu Asn

CTC AAG AAG CCT TTC ATG CCT CCT CGG GAG GTA CAT GTT CAG GTT ACC CAT TCT     218
Leu Lys Lys Pro Phe MET Pro Pro Arg Glu Val His Val Gln Val Thr His Ser

ATG CCA CCC CAA AAG ATT GAG ATC TTT AAA TCC CTA GAC AAT TGG GCT GAG GAG     272
MET Pro Pro Gln Lys Ile Glu Ile Phe Lys Ser Leu Asp Asn Trp Ala Glu Glu

AAC ATT CTG GTT CAT CTG AAG CCA GTT GAG AAA TGT TGG CAA CCG CAG GAT TTT     326
Asn Ile Leu Val His Leu Lys Pro Val Glu Lys Cys Trp Gln Pro Gln Asp Phe

TTG CCA GAT CCC GCC TCT GAT GGA TTT GAT GAG CAA GTC AGG GAA CTC AGG GAG     380
Leu Pro Asp Pro Ala Ser Asp Gly Phe Asp Glu Gln Val Arg Glu Leu Arg Glu

AGA GCA AAG GAG ATT CCT GAT GAT TAT TTT GTT GTT TTG GTT GGA GAC ATG ATA     434
Arg Ala Lys Glu Ile Pro Asp Asp Tyr Phe Val Val Leu Val Gly Asp MET Ile

ACG GAA GAA GCC CTT CCC ACT TAT CAA ACA ATG CTG AAT ACC TTG GAT GGA GTT     488
Thr Glu Glu Ala Leu Pro Thr Tyr Gln Thr MET Leu Asn Thr Leu Asp Gly Val
```

FIGURE 3B
Page 1 of 3

```
CGG GAT GAA ACA GGT GCA AGT CCT ACT TCT TGG GCA ATT TGG ACA AGG GCA TGG      542
Arg Asp Glu Thr Gly Ala Ser Pro Thr Ser Trp Ala Ile Trp Thr Arg Ala Trp

ACT GCG GAA GAG AAT AGA CAT GGT GAC CTC CTC AAT AAG TAT CTC TAC CTA TCT      596
Thr Ala Glu Glu Asn Arg His Gly Asp Leu Leu Asn Lys Tyr Leu Tyr Leu Ser

GGA CGA GTG GAC ATG AGG CAA ATT GAG AAG ACA ATT CAA TAT TTG ATT GGT TCA      650
Gly Arg Val Asp MET Arg Gln Ile Glu Lys Thr Ile Gln Tyr Leu Ile Gly Ser

GGA ATG GAT CCA CGG ACA GAA AAC AGT CCA TAC CTT GGG TTC ATC TAT ACA TCA      704
Gly MET Asp Pro Arg Thr Glu Asn Ser Pro Tyr Leu Gly Phe Ile Tyr Thr Ser

TTC CAG GAA AGG GCA ACC TTC ATT TCT CAT GGG AAC ACT GCC CGA CAA GCC AAA      758
Phe Gln Glu Arg Ala Thr Phe Ile Ser His Gly Asn Thr Ala Arg Gln Ala Lys

GAG CAT GGA GAC ATA AAG TTG GCT CAA ATA TGT GGT ACA ATT GCT GCA GAT GAG      812
Glu His Gly Asp Ile Lys Leu Ala Gln Ile Cys Gly Thr Ile Ala Ala Asp Glu

AAG CGC CAT GAG ACA GCC TAC ACA AAG ATA GTG GAA AAA CTC TTT GAG ATT GAT      866
Lys Arg His Glu Thr Ala Tyr Thr Lys Ile Val Glu Lys Leu Phe Glu Ile Asp

CCT GAT GGA ACT GTT TTG GCT TTT GCT GAT ATG ATG AGA AAG AAA ATT TCT ATG      920
Pro Asp Gly Thr Val Leu Ala Phe Ala Asp MET MET Arg Lys Lys Ile Ser MET

CCT GCA CAC TTG ATG TAT GAT GGC CGA GAT GAT AAT CTT TTT GAC CAC TTT TCA      974
Pro Ala His Leu MET Tyr Asp Gly Arg Asp Asp Asn Leu Phe Asp His Phe Ser
```

FIGURE 3B
Page 2 of 3

```
GCT GTT GCG CAG CGT CTT GGA GTC TAC ACA GCA AAG GAT TAT GCA GAT ATA TTG        1028
Ala Val Ala Gln Arg Leu Gly Val Tyr Thr Ala Lys Asp Tyr Ala Asp Ile Leu

GAG TTC TTG GTG GGC AGA TGG AAG GTG GAT AAA CTA ACG GGC CTT TCA GCT GAG        1082
Glu Phe Leu Val Gly Arg Trp Lys Val Asp Lys Leu Thr Gly Leu Ser Ala Glu

GGA CAA AAG GCT CAG GAC TAT GTT TGT CGG TTA CCT CCA AGA ATT AGA AGG CTG        1136
Gly Gln Lys Ala Gln Asp Tyr Val Cys Arg Leu Pro Pro Arg Ile Arg Arg Leu

GAA GAG AGA GCT CAA GGA AGG GCA AAG GAA GCA CCC ACC ATG CCT TTC AGC TGG        1190
Glu Glu Arg Ala Gln Gly Arg Ala Lys Glu Ala Pro Thr MET Pro Phe Ser Trp

ATT TTC GAT AGG CAA GTG AAG CTG TAGGTGGCTA AAGTGCAGGA CGAAACCGAA ATGGTTAGTT   1254
Ile Phe Asp Arg Gln Val Lys Leu

TCACTCTTTT TCATGCCCAT CCCTGCAGAA TCAGAAGTAG AGGTAGAATT TTGTAGTTGC TTTTTTATTA  1324

CAAGTCCAGT TTAGTTTAAG GTCTGTGGAA GGGAGTTAGT TGAGGAGTGA ATTTAGTAAG TTGTAGATAC  1394

AGTTGTTTCT TGTGTTGTCA TGAGTATGCT GATAGAGAGC AGCTGTAGTT TTGTTGTTGT GTTCTTTTAT  1464

ATGGTCTCTT GTATGAGTTT CTTTTCTTTC CTTTTCTTCT TTCCTTTCCT CTCTCTCTCT CTCTCTCTCT  1534

CTCTTTTTCT CTTATCCCAA GTGTCTCAAG TATAATAAGC AAACGATCCA TGTGGCAATT TTGATGATGG  1604

TGATCAGTCT CACAACTTGA TCTTTTGTCT TCTATTGGAA ACACAGCCTG CTTGTTTGAA AAAA        1668
```

FIGURE 3B
Page 3 of 3

**pCGN3235**

```
 1 TGAGAGATAGTGTGAGAGCATTAGCCTTAGAGAGAGAGAGAGAGCTTGTGTCTGAAAGAATCCACAA  69

             HindIII
             |
70 ATGGCATTGAAGCTTAACCCTTTGGCATCTCAGCCTTACAACTTCCCT 117
   METAlaLeuLysLeuAsnProLeuAlaSerGlnProTyrAsnPhePro
```

FIGURE 4A

**pCGN3236**

```
                                        PstI
                                         |
  1 ACTTCATGGGCTATTTGGACAAGAGCTTGGACTGCAGAAGAGAACCGACACGGTGATCTTCTCAATAAG  69
    ThrSerTrpAlaIleTrpThrArgAlaTrpThrAlaGluGluAsnArgHisGlyAspLeuLeuAsnLys

 70 TATCTTTACTTGTCTGGACGTGTTGACATGAGGCAGATTGAAAAGACCATTCAGTACTTGATTGGTTCT 138
    TyrLeuTyrLeuSerGlyArgValAspMETArgGlnIleGluLysThrIleGlnTyrLeuIleGlySer

          BamHI
           |
139 GGAATGGATCCTAGAACAGAGAACAATCCTTACCTCGG 176
    GlyMETAspProArgThrGluAsnAsnProTyrLeuAla
```

FIGURE 4B

**pCGN3235**

```
TGAGAGATAG TGTGAGAGCA TTAGCCTTAG AGAGAGAGAG AGAGAGCTTG TGTCTGAAAG AATCCACAA

ATG GCA TTG AAG CTT AAC CCT TTG GCA TCT CAG CCT TAC AAC TTC CCT TCC TCG
MET Ala Leu Lys Leu Asn Pro Leu Ala Ser Gln Pro Tyr Asn Phe Pro Ser Ser

GCT CGT CCG CCA ATC TCT ACT TTC AGA TCT CCC AAG TTC CTC TGC CTC GCT TCT
Ala Arg Pro Pro Ile Ser Thr Phe Arg Ser Pro Lys Phe Leu Cys Leu Ala Ser

TCT TCT CCC GCT CTC AGC TCC AAG GAG GTT GAG AGT TTG AAG AAG CCA TTC ACA
Ser Ser Pro Ala Leu Ser Ser Lys Glu Val Glu Ser Leu Lys Lys Pro Phe Thr

CCA CCT AAG GAA GTG CAC GTT CAA GTC CTG CAT TCC ATG CCA CCC CAG AAG ATC
Pro Pro Lys Glu Val His Val Gln Val Leu His Ser MET Pro Pro Gln Lys Ile

GAG ATC TTC AAA TCC ATG GAA GAC TGG GCC GAG CAG AAC CTT CTA ACT CAG CTC
Glu Ile Phe Lys Ser MET Glu Asp Trp Ala Glu Gln Asn Leu Leu Thr Gln Leu

AAA GAC GTG GAG AAG TCG TGG CAG CCC CAG GAC TTC TTA CCC GAC CCT GCA TCC
Lys Asp Val Glu Lys Ser Trp Gln Pro Gln Asp Phe Leu Pro Asp Pro Ala Ser

GAT GGG TTC GAA GAT CAG GTT AGA GAG CTA AGA GAG AGG GCA AGA GAG CTC CCT
Asp Gly Phe Glu Asp Gln Val Arg Glu Leu Arg Glu Arg Ala Arg Glu Leu Pro

GAT GAT TAC TTC GTT GTT CTG GTG GGA GAC ATG ATC ACG GAA GAG GCG CTT CCG
Asp Asp Tyr Phe Val Val Leu Val Gly Asp MET Ile Thr Glu Glu Ala Leu Pro
```

12/42

FIGURE 4C
Page 1 of 3

```
ACC TAT CAA ACC ATG TTG AAC ACT TTG GAT GGA GTG AGG GAT GAA ACT GGC GCT
Thr Tyr Gln Thr MET Leu Asn Thr Leu Asp Gly Val Arg Asp Glu Thr Gly Ala

AGC CCC ACT TCA TGG GCT ATT TGG ACA AGA GCT TGG ACT GCA GAA GAG AAC CGA
Ser Pro Thr Ser Trp Ala Ile Trp Thr Arg Ala Trp Thr Ala Glu Glu Asn Arg

CAC GGT GAT CTT CTC AAT AAG TAT CTT TAC TTG TCT GGA CGT GTT GAC ATG AGG
His Gly Asp Leu Leu Asn Lys Tyr Leu Tyr Leu Ser Gly Arg Val Asp MET Arg

CAG ATT GAA AAG ACC ATT CAG TAC TTG ATT GGT TCT GGA ATG GAT CCT AGA ACA
Gln Ile Glu Lys Thr Ile Gln Tyr Leu Ile Gly Ser Gly MET Asp Pro Arg Thr

GAG AAC AAT CCT TAC CTC GGC TTC ATC TAC ACT TCA TTC CAA GAA AGA GCC ACC
Glu Asn Asn Pro Tyr Leu Gly Phe Ile Tyr Thr Ser Phe Gln Glu Arg Ala Thr

TTC ATC TCT CAC GGA AAC ACA GCT CGC CAA GCC AAA GAG CAC GGA GAC CTC AAG
Phe Ile Ser His Gly Asn Thr Ala Arg Gln Ala Lys Glu His Gly Asp Leu Lys

CTA GCC CAA ATC TGC GGC ACA ATA GCT GCA GAC GAG AAG CGT CAT GAG ACA GCT
Leu Ala Gln Ile Cys Gly Thr Ile Ala Ala Asp Glu Lys Arg His Glu Thr Ala

TAC ACC AAG ATA GTT GAG AAG CTC TTT GAG ATT GAT CCT GAT GGT ACT GTG ATG
Tyr Thr Lys Ile Val Glu Lys Leu Phe Glu Ile Asp Pro Asp Gly Thr Val MET

GCG TTT GCA GAC ATG ATG AGG AAG AAA ATC TCG ATG CCT GCT CAC TTG ATG TAC
Ala Phe Ala Asp MET MET Arg Lys Lys Ile Ser MET Pro Ala His Leu MET Tyr
```

FIGURE 4C
Page 2 of 3

```
GAT GGG CGG GAT GAA AGC CTC TTT GAC AAC TTC TCT TCT GTT GCT CAG AGG CTC
Asp Gly Arg Asp Glu Ser Leu Phe Asp Asn Phe Ser Ser Val Ala Gln Arg Leu

GGT GTT TAC ACT GCC AAA GAC TAT GCG GAC ATT CTT GAG TTT TTG GTT GGG AGG
Gly Val Tyr Thr Ala Lys Asp Tyr Ala Asp Ile Leu Glu Phe Leu Val Gly Arg

TGG AAG ATT GAG AGC TTG ACC GGG CTT TCA GGT GAA GGA AAC AAA GCG CAA GAG
Trp Lys Ile Glu Ser Leu Thr Gly Leu Ser Gly Glu Gly Asn Lys Ala Gln Glu

TAC TTG TGT GGG TTG ACT CCA AGA ATC AGG AGG TTG GAT GAG AGA GCT CAA GCA
Tyr Leu Cys Gly Leu Thr Pro Arg Ile Arg Arg Leu Asp Glu Arg Ala Gln Ala

AGA GCC AAG AAA GGA CCC AAG GTT CCT TTC AGC TGG ATA CAT GAC AGA GAA GTG
Arg Ala Lys Lys Gly Pro Lys Val Pro Phe Ser Trp Ile His Asp Arg Glu Val

CAG CTC TAA AAAGGAA CAAAGCTATG AAACCTTTTC ACTCTCCGTC GTCCCTCATT TGATCTATCT
Gln Leu  *

GCTCTTGAAA TTGGTGTAGA TTACTATGGT TTGTGATATT GTTCGTGGGT CTAGTTACAA AGTTGAGAAG

CAGTGATTTA GTAGCTTTGT TGTTTCCAGT CTTTAAATGT TTTTGTGTTT GGTCCTTTTA GTAAACTTGT

TGTAGTTAAA TCAGTTGAAC TGTTTGGTCT GT
```

FIGURE 4C
Page 3 of 3

```
GAT GCC AAA ANG CCT CAC ATG CCT CCT AGA GAA GCT CAT GTG CAA AAG    48
Asp Ala Lys Xaa Pro His MET Pro Pro Arg Glu Ala His Val Gln Lys
 1               5                  10                  15

ACC CAT TCA ATK CCG CCT CAA AAG ATT GAG ATT TTC AAA TCC TTG GAG    96
Thr His Ser Xaa Pro Pro Gln Lys Ile Glu Ile Phe Lys Ser Leu Glu
            20                  25                  30

GGT TGG GCT GAG GAG AAT GTC TTG GTG CAT CTT AAA CCT GTG GAG AA    143
Gly Trp Ala Glu Glu Asn Val Leu Val His Leu Lys Pro Val Glu
        35                  40                  45
```

FIGURE 5

**Amino Acid Sequence From Fragment F2**

K E I P D D Y FVVLVGDMITEEALPTY Q T M L N T

AAA GAA AUU CCN GAU GAU UAU CAA ACN AUG CUN AAU AC/N
G G C C C G C
A

Forward Primers:

5'GCTAAGCTT AAP GAP ATQ CCA GAQ GAQ TA3' Desat 13-1
A CCG Desat 13-2
CCC Desat 13-3
CCT Desat 13-4

Reverse Primers:
(complements)

Desat 13-5a 3' GTQ TGN TAC GAN TTP TGCTTAAGCGA 5'
Desat 13-6a AAQ

Oligonucleotides

P = A or G
Q = T or C
N = A, C, T or C

16/42

FIGURE 6

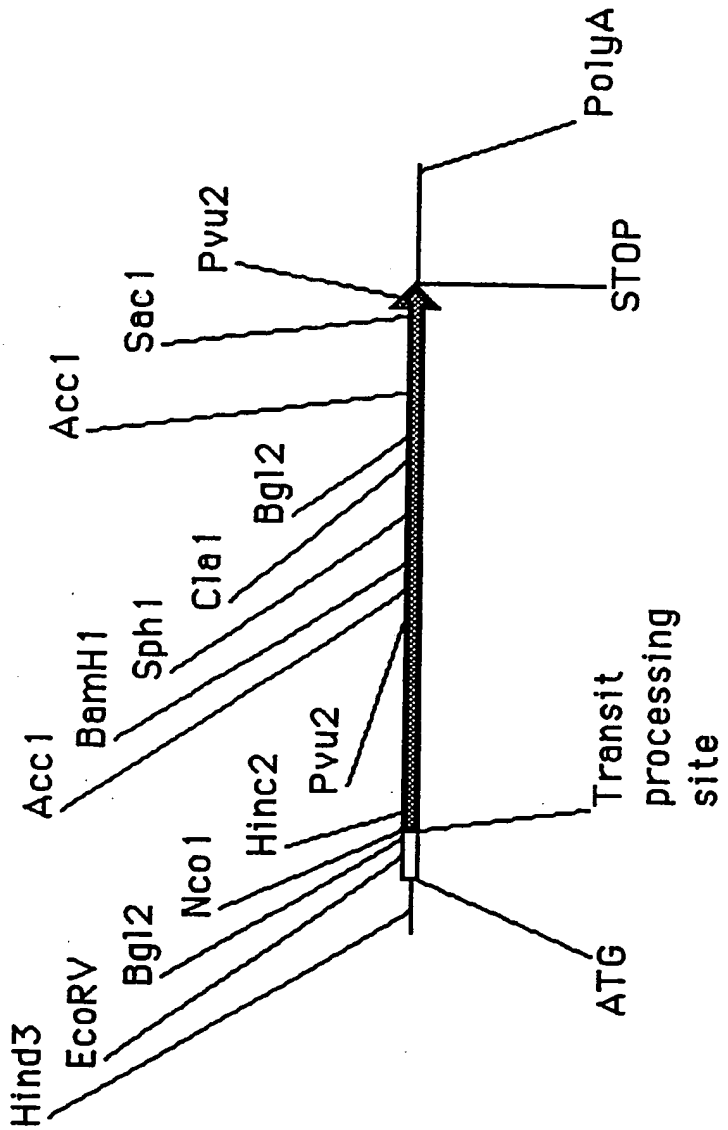

PolyA
STOP
Pvu2
Sac1
Acc1
Bgl2
Cla1
Sph1
BamH1
Acc1
Pvu2
Hinc2
Nco1
Bgl2
EcoRV
Hind3
Transit processing site
ATG

FIGURE 7A

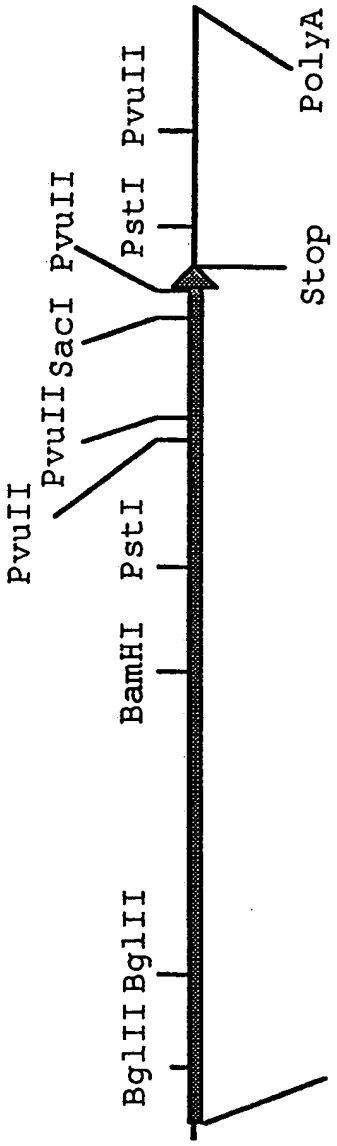

PolyA
PvuII
PstI
PvuII
Stop
PvuII
SacI
PvuII
PvuII
PstI
BamHI
BglII
BglII
ATG

FIGURE 7B

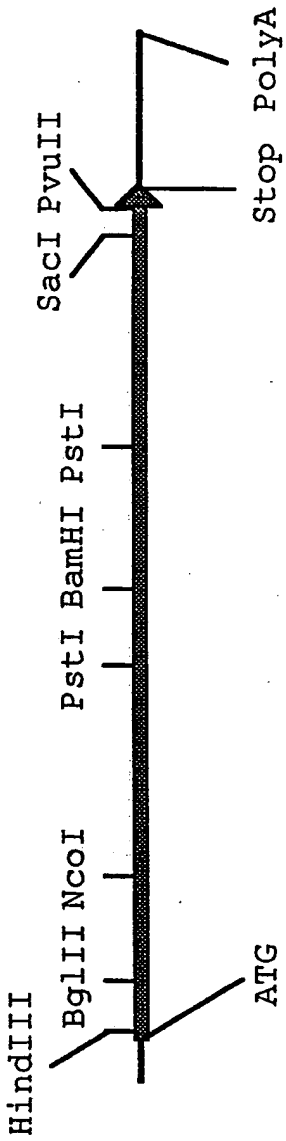

HindIII
BglII
NcoI
PstI
BamHI
PstI
SacI
PvuII
ATG
Stop
PolyA

FIGURE 7C

```
TCTAGAATTC TCTAATTACG TCTGTTTGTT CTATTTTTTA TATGATATCA AATATTCGTC ATAAATATAT    70
GGTTTAAGAT GCCAAAAAAT TATTTACTTG GTGAATATAA TACGTTAAAT ATTAGAAATA CATCATTTAG   140
TTAAATAAAT AACCAAAAAC CAAAAATTCA TATCCGCGCT GGCGCGCGGT CAGGGTCTCG TTAGTTTTAA   210
AATCAATGCA GTTTACAATT AATTTCCAGC TGAAAATAAG TATAATTTGT ATTGAAATTA TAAAGTGACA   280
TTTTTTGTGT AACAAATATT TTGTGTAACA AGAATTAAAA AAAAAAACAG AAAATACTCA GCTTTTTTAA   350
TAATAAAAAA AATTAATTGA GTTAGAAAAT TGTTGTACCA ATAACAAAAG ATTTATATGG AATTATAAAA   420
TCAACACACC AATAACACAA GACTTTTTAA AAATTTAAGA ATAATATAAG CAATAACAAT AGAATCTTCA   490
AATTCTTCAA ATCCTTAAAA ATCAATCTCC CACTATTAAT CCCCCTTAGT TTTAGTTGGT AATGGCAACG   560
TTTGTTGACT ACCGTATTGT AACTTTTGTC AAATTGTCAT AAATACGTGT CAAACTCTGG TAAAAAATTA   630
GTCTGCTACA TCTGTCTTTT ATTTATAAAA CACAGCTGTT AATCAGAATT TGGTTTATTA AATCAACAAC   700
CTGCACGAAA CTTGTGTGAG CATATTTTGT CTGTTTCTGG TTCATGACCT TCTTCCGCAT GATGGCCAAG   770
TGTAATGGCC ACTTGCAAGA GCGTTTCTTC AACGAGATAA GTCGAACAAA TATTTGTCCG TTACGACCAC   840
ATATAAAATC TCCCCATCTC TATATATAAT ACCAGCATTC ACCATCATGA ATACCTCAAA TCCCAATCTC   910
ACAAATACTT CAATAAAAAG ACCAAAAAAA ATTAAAGCAA AGAAAAGCCT TCTTGTGCAC AAAAAAAAAA   980
```

FIGURE 8
Page 1 of 4

```
GAAGCCTTCT AGGTTTTCAC GAC ATG AAG TTC ACT ACT CTA ATG GTC ATC ACA TTG          1036
                          MET Lys Phe Thr Thr Leu MET Val Ile Thr Leu

GTG ATA ATC GCC ATC TCG TCT CCT GTT CCA ATT AGA GCA ACC ACG GTT GAA AGT       1090
Val Ile Ile Ala Ile Ser Ser Pro Val Pro Ile Arg Ala Thr Thr Val Glu Ser

TTC GGA GAA GTG GCA CAA TCG TGT GTT GTG ACA GAA CTC GCC CCA TGC TTA CCA       1144
Phe Gly Glu Val Ala Gln Ser Cys Val Val Thr Glu Leu Ala Pro Cys Leu Pro

GCA ATG ACC ACG GCA GGA GAC CCG ACT ACA GAA TGC TGC GAC AAA CTG GTA GAG       1198
Ala MET Thr Thr Ala Gly Asp Pro Thr Thr Glu Cys Cys Asp Lys Leu Val Glu

CAG AAA CCA TGT CTT TGT GGT TAT ATT CGA AAC CCA GCC TAT AGT ATG TAT GTT       1252
Gln Lys Pro Cys Leu Cys Gly Tyr Ile Arg Asn Pro Ala Tyr Ser MET Tyr Val

ACT TCT CCA AAC GGT CGC AAA GTC TTA GAT TTT TGT AAG GTT CCT TTT CCT AGT       1306
Thr Ser Pro Asn Gly Arg Lys Val Leu Asp Phe Cys Lys Val Pro Phe Pro Ser

TGT TAAATCTCTC AAGACATTGC TAAGAAAAAT ATTATTAAAA ATAAAAGAAT CAAACTAGAT       1369
Cys

CTGATGTAAC AATGAATCAT CATGTTATGG TTGAAGCTTA TATGCTGAAG TGTTTGATTT TATATATGTG  1439

TGTGTGTGTG TCCTGCTCAA TTTTTGAAAC ACACACGTTT CTCCTGATTT GGATTTAAAT TATATTTTGA  1509

GTTAAAAAAA AGAAAAAGAT GGAATGCTAT TTATACAAGT TGATGAAAAA GTGGAAGTAC AATTTAGATA  1579
```

FIGURE 8
Page 2 of 4

```
TCTCCTACAC TTAAAGAATG AAACAATAAT AGACTTACGA AACAAATGAA AAATACATAA ATTGTCGACA  1649
ATCAACGTCC GATGACGAGT TTATTATTAA AAATTTGTGT GAAGGACTAG CAGTTCAACC AAATGATATT  1719
GAACATATAC ATCAACAAAT ATGATAATCA TAAAAGAGAG AATGGGGGGG GGGTGTCGTT TACCAGAAAC  1789
CTCTTTTTCT CAGCTCGCTA AAACCCTACC ACTAGAGACC TAGCTCTGAC CGTCGGCTCA TCGGTGCCGG  1859
AGGTGTAACC TTTCTTTCCC ATGACCCGAA ACCTCTCTTT CCCAACTCAC GAAAACCCTA CAATCAAAAA  1929
CCTAGCTCCG ACCGTCGGCT CATCGGTGCC GAAGGTGTAA CCTTTCTCTC CCATCATAGT TTCTCGTAAA  1999
TGAAAGCTAA TTGGGCAATC GATTTTTTAA TGTTTAAACC ATGCCAAGCC ATTTCTTATA GGACAATTGT  2069
CAATAATAGC ATCTTTTGAG TTTTGTCTCA AAAGTGACAC TAGAAGAAAA AAGTCACAAA AATGACATTC  2139
ATTAAAAAGT AAAATATCCC TAATACCTTT GGTTTAAATT AAATAAGTAA ACAAAAATAA ATAAAAACAA  2209
ATAAAATAAA AATAAAAAAT GAAAAAAAGA AATTTTTTTA TAGTTTCAGA TTATATGTTT TCAGATTCGA  2279
AATTTTTTAA ATTCCCTTTT TTAAATTTTC TTTTTTGAAA TTTTTTTTTT TGAAATTTTT TGAAACTGTT  2349
TTTAAAATTT TTATTTTTAA TTTTTTAGTA TTTATTTTTT ATTTTATAAA ATTTTAAACG CTAATTCCAA  2419
AACTCCCCCC CCCCCCCCCC CCCCAATTCT CTCCTAGTCT TTTTCTCTTT CTTATATTTG GGCTTCTATC  2489
TTCTCTTTTT TTTTCAGGCC CAAAGTATCA TGTGTAACAA CCGGTGTTCA AAAACGCGCC CGCCTGGCCG  2559
```

FIGURE 8
Page 3 of 4

```
TTTACTCGCC CGATTAAATG ATGATCGGAA GGCTGCCATG GCGAGGCGGA GGTAATCAGT GGTTCTAGGC  2629
GCTGAAACTA GAAAACCTTC AAAAATCGAA ATTTTAAGAG CTAAATCGGT GTTTATCTCA TGAATCTATT  2699
ATATTTAGTT GAAACTCACA AGAATCGGTT GTAAAAACTA TGAAATCGTG CAAAAAAAAT GAAGAACAAA  2769
ATATTCTCAG ATCTGGAAAA CACAGAGAAG AGGTTGAAGA TGAGGGTAAA ATCGTATTTT GTCATTCATT  2839
AAACTAAAAT CAAAAAAAAA TGATGCAAAA TTCAATGATA ATAACTCGAA CTCGCAACCA TATGCATCTT  2909
TAGACTGCGA CACGGACCAC TAGACTAAGC AATTTTAATG TTTATTCATC ACAGACCTAA TATATGTCTA  2979
AAACTAGGCG CCGAGTACGC CCCGCTTAAT CCCGAGTTTT TGTTAGCTCG CTAGACCCAG GGTCACCGCC  3049
CGACTAACGA GTAGCGTAAT TCTGAACTGG GGTAACAACA TAGAGAACAT CGCCGACCCT TCCCTGCCGA  3119
TGATGCCGCC TCCGATGAAC TTCCTGTAAC GCCTTCAGTT TCCATTGATT TTCCCCTTTA ATCTGATCAG  3189
TTCCATGTTT TATCCAACTC ATCCCACTCC GTAGCATTTA ATCGATCTCA TCATTTACAT ACATAACCAG  3259
TAGGAGGTCT CATATAAATT TGAACGTTTC CAGCGATGAA CAGTGCCAAT CTCTGCGAAA TCCATTTCTC  3329
TAAGCTCAGG GCTGGCGGCT GCAGCCCGGG GATCCACTAG TTCTAGGCGG CCGCACCGCG GTGGAGCTCC  3399
AATTCGCCCT ATAGTGAGTC GTATTACGCG CGCTCACTGG C                                  3440
```

FIGURE 8
Page 4 of 4

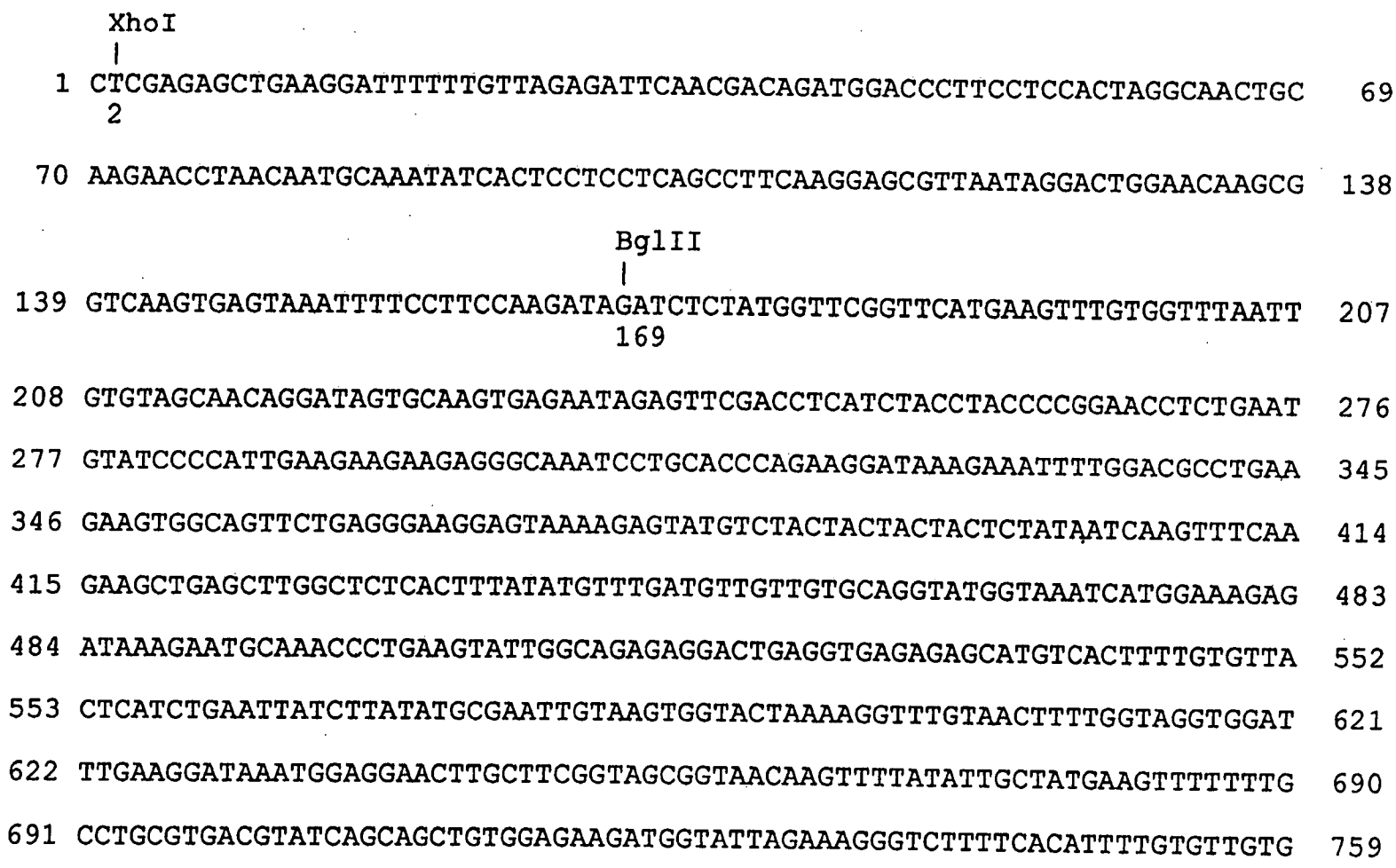



FIGURE 9
Page 1 of 6

```
760 ACAAATATTAATTCGGCCGGTATGGTTTGGTTAAGACTTGTTGAGAGACGTGTGGGGTTTTTTGATGTA  828

829 TAATTAGTCTGTGTTTAGAACGAAACAAGACTTGTTGCGTATGCTTTTTTTAACTTGAGGGGGTTTGTT  897

                                          BglII
                                            |
898 GTTGTTAGTTAGGAACTTGACTTTGTCTCTTTCTCTCAAGATCTGATTGGTAAGGTCTGGGTGGTAGTA  966
                                          937

967 CTGTTTGGTTTAATTTGTTTTGACTATTGAGTCACTGTGGCCCATTGACTTTAAATTAGGCTGGTATAT 1035

1036 TTTTTGGTTTAAAACCGGTCTGAGATAGTGCAATTTCGATTCAGTCAATTTTAAATTCTTCAAGGTAAT 1104

1105 GGGCTGAATACTTGTATAGTTTTAAGACTTAACAGGCCTTAAAAGGCCCATGTTATCATAAAACGTCAT 1173

                    HindIII
                       |
1174 TGTTTAGAGTGCACCAAGCTTATAAAATGTAGCCAGGCCTTAAAAGACTTAACAGGCCTTAAAAGACTT 1242
                    1190

1243 AACATTCCTTAAAAGGCCCATGTTATCATAAAACGTCATCGTTTTGAGTGCACCAAGCTAAATGTAGCC 1311

1312 AGGCCTTAAAAGACTTAACAGGCCTTAAAAGGCCCATGTTATCATAAAACGCCGTCGTTTTGAGTGCAC 1380
```

FIGURE 9
Page 2 of 6

```
        HindIII
          |
1381 CAAGCTTATAAATGTAGCCAGCTACCTCGGGACATCACGCTCTTTGTACACTCCGCCATCTCTCTCTCT 1449
          1383

        XhoI    BglII                       SalI
         |       |                          |
1450 CTCGAGCAGATCTCTCTCGGGAATATCGACAATGTCGACCACTTTCTGCTCTTCCGTCTCCATGCAAGC 1518
          1451   1458                       1484

1519 CACTTCTCTGGTAATCTCATCTCCTTCTTGTGTTCCCAGATCGCTCTGATCATACTTTCTTTTAGATCA 1587

1588 TTTGCCTCTGATCTGTTGCTTGATGTTTGTTAACTCTCCACGCATGTTTGATTATGTTGAGAATTAGAA 1656

1657 AAAAAATGTTAGCTTTACGAATCTTTAGTGATCATTTCAATTGGATTTGCAATCTTGTGTGACATTTGA 1725

1726 GGCTTGTGTAGATTTCGATCTGTATTCATTTTGAATCACAGCTATAATAGTCATTTGAGTAGTAGTGTT 1794

1795 TTTAAATGAACATGTTTTGTTGTATTGATGGAACAAACAGGCAGCAACAACGAGGATTAGTTTCCAGAA 1863

1864 GCCAGCTTTGGTTTCAACGACTAATCTCTCCTTCAACCTCCGCCGTTCAATCCCCACTCGTTTCTCAAT 1932

1933 CTCCTGCGCGGTATGTTCTCATTCTCAGCATTTATTTCGAGCTTGCTTGTCATGGTACTCTCTCTAATT 2001

2002 GTCTATTTGGTTTATTAGGCCAAACCAGAGACGGTTGAGAAAGTGTCTAAGATAGTTAAGAAGCAGCTA 2070
```

FIGURE 9
Page 3 of 6

```
2071 TCACTCAAAGACGACCAAAAGGTCGTTGCGGAGACCAAGTTTGCTGATCTTGGAGCAGATTCTCTCGAC 2139

2140 ACTGTAAGTCATCAATCATTCTCTTATGTGAATAAAGAGAACTTGAAGAGTTTGTTTTTAACATATTAA 2208

                                                          EcoRV
                                                            |
2209 CTGAGTGTTTTGCATGCAGGTTGAGATAGTGATGGGTTTAGAGGAAGAGTTTGATATCGAAATGGCTGA 2277
                                                          2264

                                                             SstI
                                                              |
2278 AGAGAAAGCTCAGAAGATTGCTACTGTGGAGGAAGCTGCTGAACTCATTGAAGAGCTCGTTCAACTTAA 2346
                                                             2335

2347 GAAGTAATTTTAGTATTAAGAGCAGCCAAGGCTTTGTTGGGTTTGTTGTTTTCATAATCTTCCTGTCAT 2415

2416 TTTCTTTTTCTTTAATGTGTCAAGCGACTCTGTTGGTTTAAAGTAGTATCTGTTTGCCATGGATCTCTC 2484

          SalI                                HindIII
           |                                    |
2485 TCTATTTGTCGACTGAAAACTTTTGGTTTACACATGAAAGCTTGTTCTTGTTCTTTCTTAAATCGAAAT 2553
          2493                                2523

2554 GCCAAATGCGAGATTAGGGAATCTTGTATTAACACATACATAAGTCAAAGAGTAGGCCCTAAGATGACA 2622

2623 ATTTATAAACAATCCTATTCACATTGTATATACAGGTTATGATTATTCCCAATCAGCGTCAAAGAATCC 2691
```

FIGURE 9
Page 4 of 6

```
2692 AGCATCTTTCATCTCTGAATAGTAGACATTCTCCAAGTTCACATCTTCCTCCTGCACCAAAAACCAGTA 2760
2761 CTAAATCATGAACATTGCAATAATCACATGCCTAGGCGAGAGTTTTGGTGATGTGGTGTTAGTGATAGT 2829
2830 GATACTGATGGTGCTAGAGCGGTTAAGAAGGATTAACCTGGAAGAAGTCTGCAAGGAAAGTAACATAGA 2898
2899 GAAGAGGAAGATAGGAGTGGTAACAAACACTTGTGATCCCATACAGCCTCCCAGCATTTTTCAAATGTT 2967
2968 ATTTCCTTACATAAAGAAACAAGAGAAGTCTGACTAGATGATATTTATATAGGATAAGTGTTTTACCAT 3036
3037 AAGCCAAAGTGAGCGCCGTTTGCAAGAGCTAACCAGACAGTACACGTTTGGCATATATCTCATCAACAT 3105
3106 GATCTGAAAAGTAACATATCACAGTTAATGAACACAATGGTTACCTTGAGAAGCAAATCAAGACCTATA 3174
3175 ACAAGCCCAGAGATGAGGAAAGTCCGTGTCAACGCTTCACCGCCATTCGCGTAGTTTCCTTGGAAGACA 3243
3244 AAGGCCACCAACCAAACTTACTTCCAGAAACAACACTCCAAATGTTGTCAACAAAGTCAATAGATTCCA 3312
3313 AACTACTTCGTTACAGGGTTGTATAGATAATATAATAGAATAGTGGGAAGATAGTATAAATAAAATAAA 3381
3382 TAAAAGATCCTATCGGTAAATAGTTTATAATATCGGGGGCGTATATAAAGTATAAAAGAAACTCTTCTC 3450
3451 CAATCCGACCGTTGAAAATCACTCTCAATCTCTGGCGTAACGACCGGATCGTTCGCGCGTAATTTTCGC 3519
3520 TGCTATAAATAGAAACTTTCCTCTTCTGTTTCTCGATCAAAATTTTTTTTTGGAAAAATTAAGTTTGAA 3588
```

28/42

FIGURE 9
Page 5 of 6

```
3589 TCTATCGTAGATGCTGTGACAAAAAAAAATTGTTTTATCGAAGATGAGAAACATGAGGCCTGTTCATGC 3657

                     BamHI
                     |
3658 AAGGAACCAGACCACGGATCCATCTTCGCCGATGATGACGTCTCCTCTGATGAATCGTCACGCACGGAC 3726
                     3674

       BamHI
       |
3727 AGGATCCAACGCTGGACCAGCATCTAACGCCAAGAAAGCACAGACGAAAGCAGCAGCTCAGAGACTCGC 3795
       3729

3796 GGCTGTGATGTCGAACCAAACAGGCGACGATGAAGACAGTGATGATGACCTTTCCTTTGACTACAACGC 3864

                                    BglII
                                    |
3865 TGTCGGAAGCATTGGTCTCGCTGCCGGAAGATCT 3898
                                    3894
```

FIGURE 9
Page 6 of 6

Lambda CGN1-2

NCG-186 Linear LENGTH = 4325

```
                                        SduI
     XhoI                              NlaIV
     AvaI                                 HgiJII      HindIII
      |                                   |  |          |
   1 CTCGAGGCAGTCACTAACATGAAGTTTGACGAGGAGCCCAACTATGGGAAGCTTATTTCTCTTTTCGAT   69
     2                                      39         50
                                          36

                                 XbaI                          SacI
                                  |                             |
  70 ACTCTAATTGAGCCGTGCGCTCTATCTAGACCAATTAGAATTGATGGAGCTCTAAAGGTTGCTGGCTGT  138
                                 95                            121

                  NdeI                          SspI                          NdeI
                   |                             |                             |
 139 TTTCTTGTTCATATGATTAACTTCTAAACTTGTGTATAAATATTCTCTGAAAGTGCTTCTTTTGGCATA  207
                  150                           180                           206

                                        Ksp632I
                                           |
 208 TGTAGGTTGGGCAAAAACGAGGAAGATTGCTTCTCAATTTGGAAGAGGATGAACAGCCGAAGAAGAAAA  276
                                          245
```

FIGURE 10
Page 1 of 13

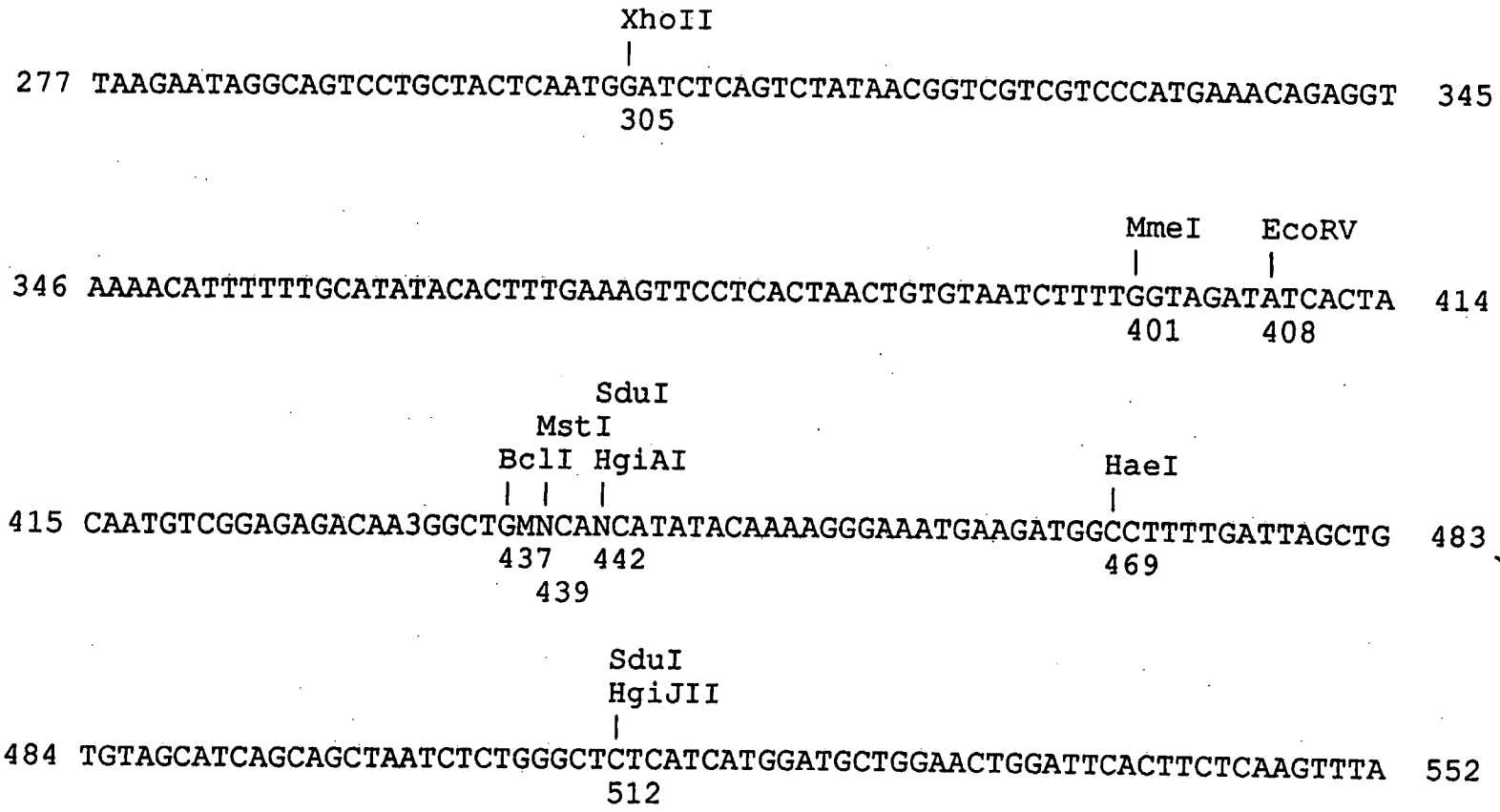



FIGURE 10
Page 2 of 13

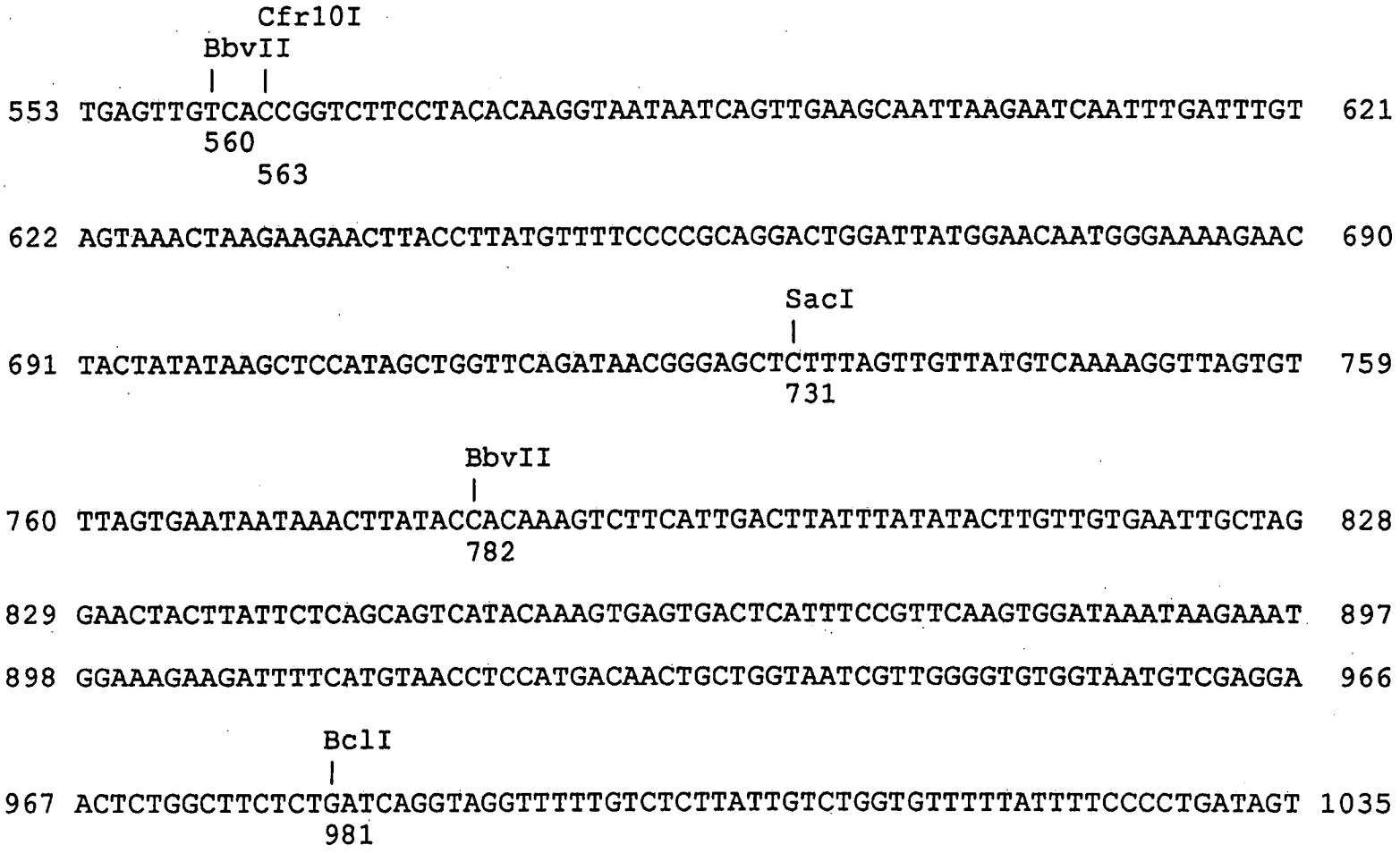



FIGURE 10
Page 3 of 13


FIGURE 10
Page 4 of 13

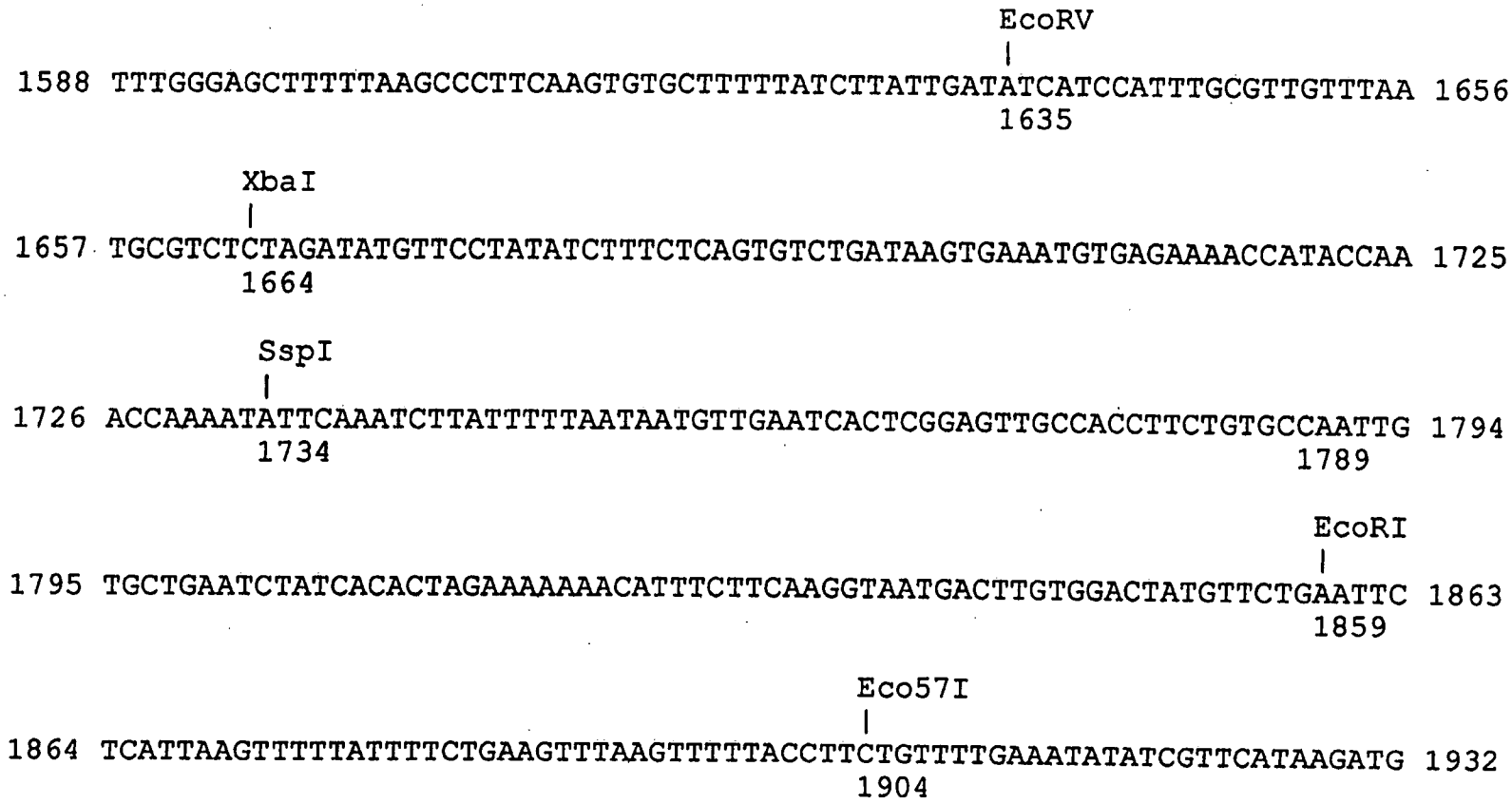
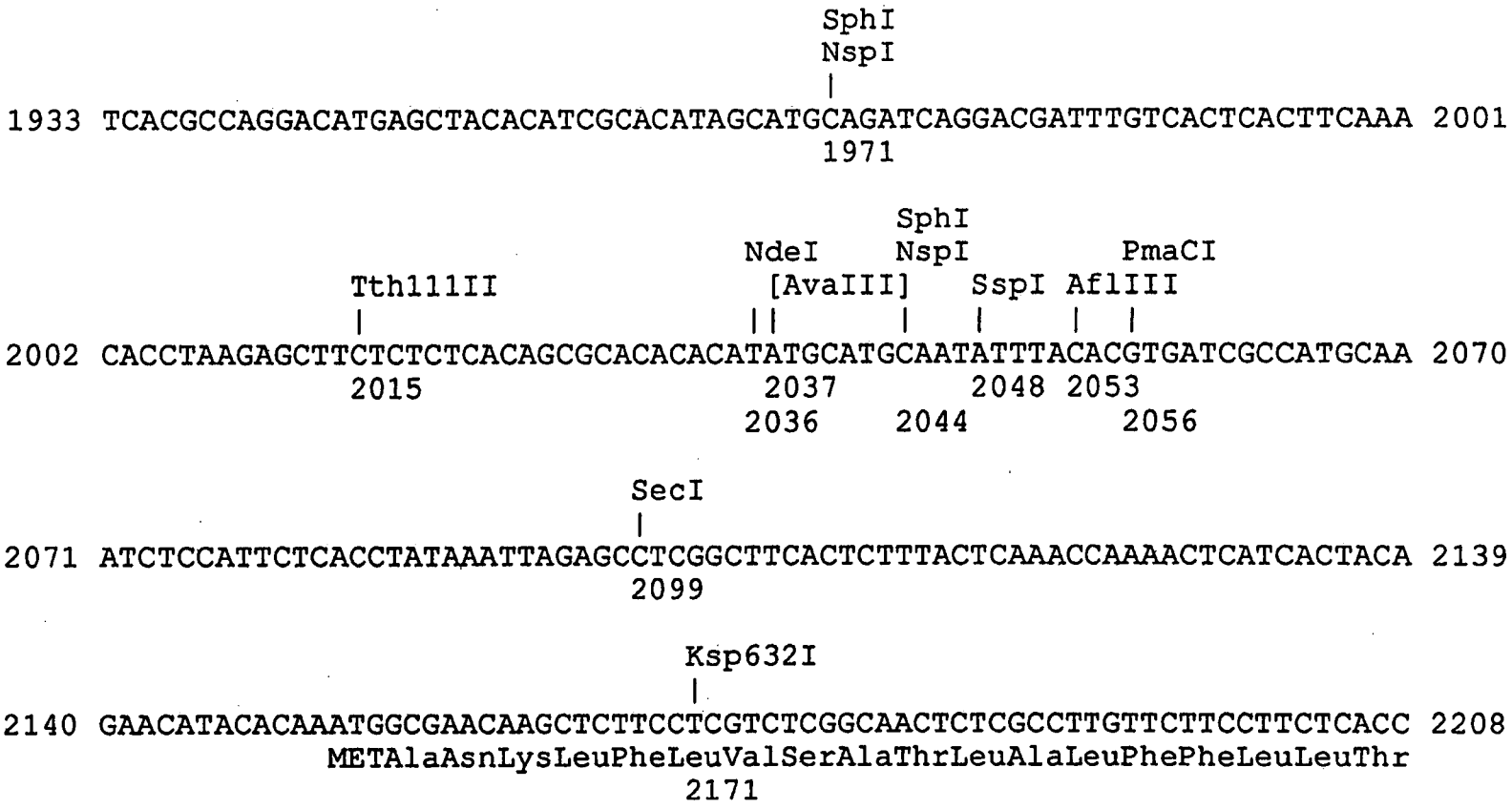



FIGURE 10
Page 5 of 13




35/4/2

FIGURE 10
Page 6 of 13

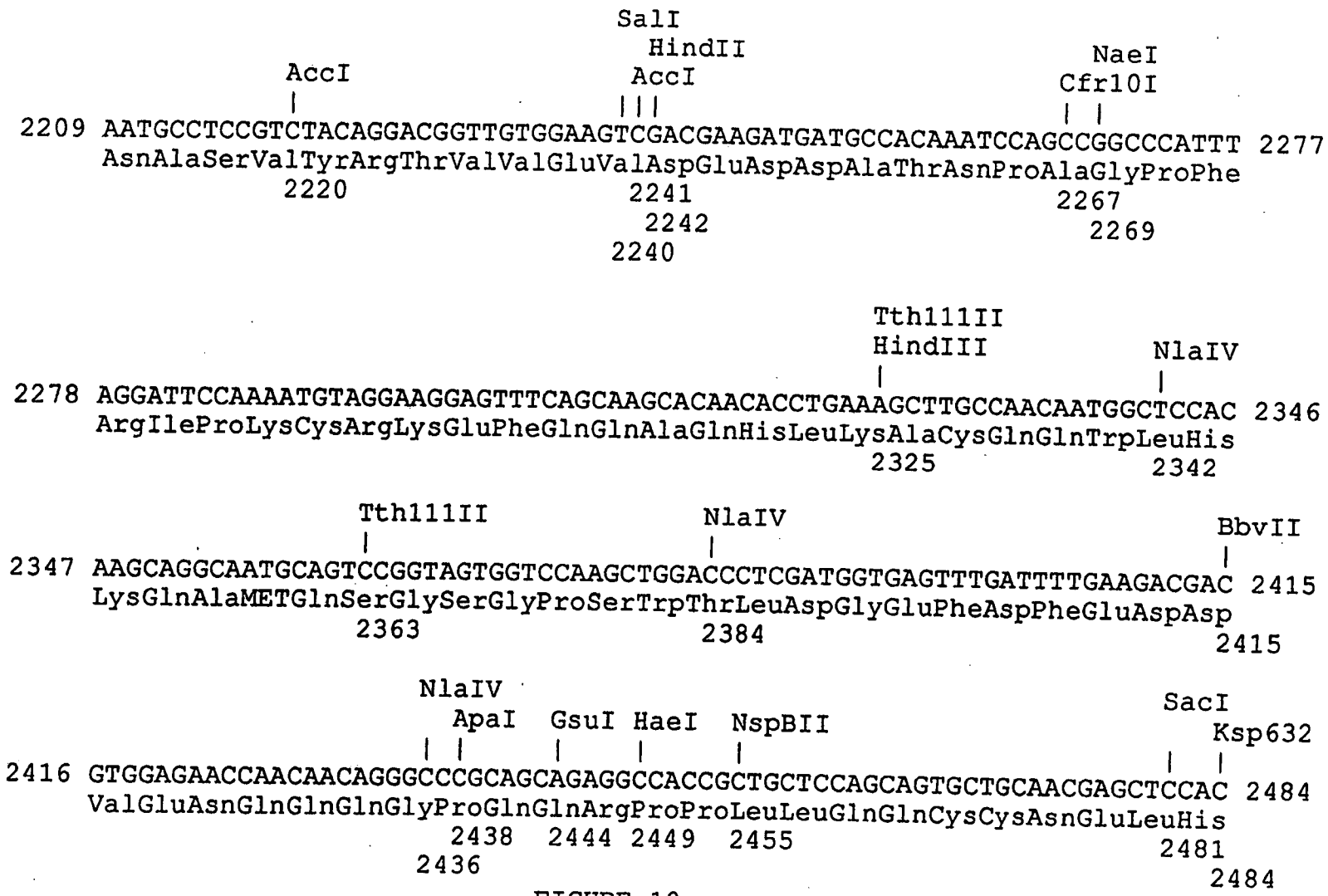

SalI
HindII
AccI
AccI
NaeI
Cfr10I
2209 AATGCCTCCGTCTACAGGACGGTTGTGGAAGTCGACGAAGATGATGCCACAAATCCAGCCGGCCCATTT 2277
AsnAlaSerValTyrArgThrValValGluValAspGluAspAspAlaThrAsnProAlaGlyProPhe
2220
2241
2242
2240
2267
2269
Tth111II
HindIII
NlaIV
2278 AGGATTCCAAAATGTAGGAAGGAGTTTCAGCAAGCACAACACCTGAAAGCTTGCCAACAATGGCTCCAC 2346
ArgIleProLysCysArgLysGluPheGlnGlnAlaGlnHisLeuLysAlaCysGlnGlnTrpLeuHis
2325
2342
Tth111II
NlaIV
BbvII
2347 AAGCAGGCAATGCAGTCCGGTAGTGGTCCAAGCTGGACCCTCGATGGTGAGTTTGATTTTGAAGACGAC 2415
LysGlnAlaMETGlnSerGlySerGlyProSerTrpThrLeuAspGlyGluPheAspPheGluAspAsp
2363
2384
2415
NlaIV
ApaI
GsuI
HaeI
NspBII
SacI
Ksp632
2416 GTGGAGAACCAACAACAGGGCCCGCAGCAGAGGCCACCGCTGCTCCAGCAGTGCTGCAACGAGCTCCAC 2484
ValGluAsnGlnGlnGlnGlyProGlnGlnArgProProLeuLeuGlnGlnCysCysAsnGluLeuHis
2438
2444
2449
2455
2436
2481
2484

FIGURE 10

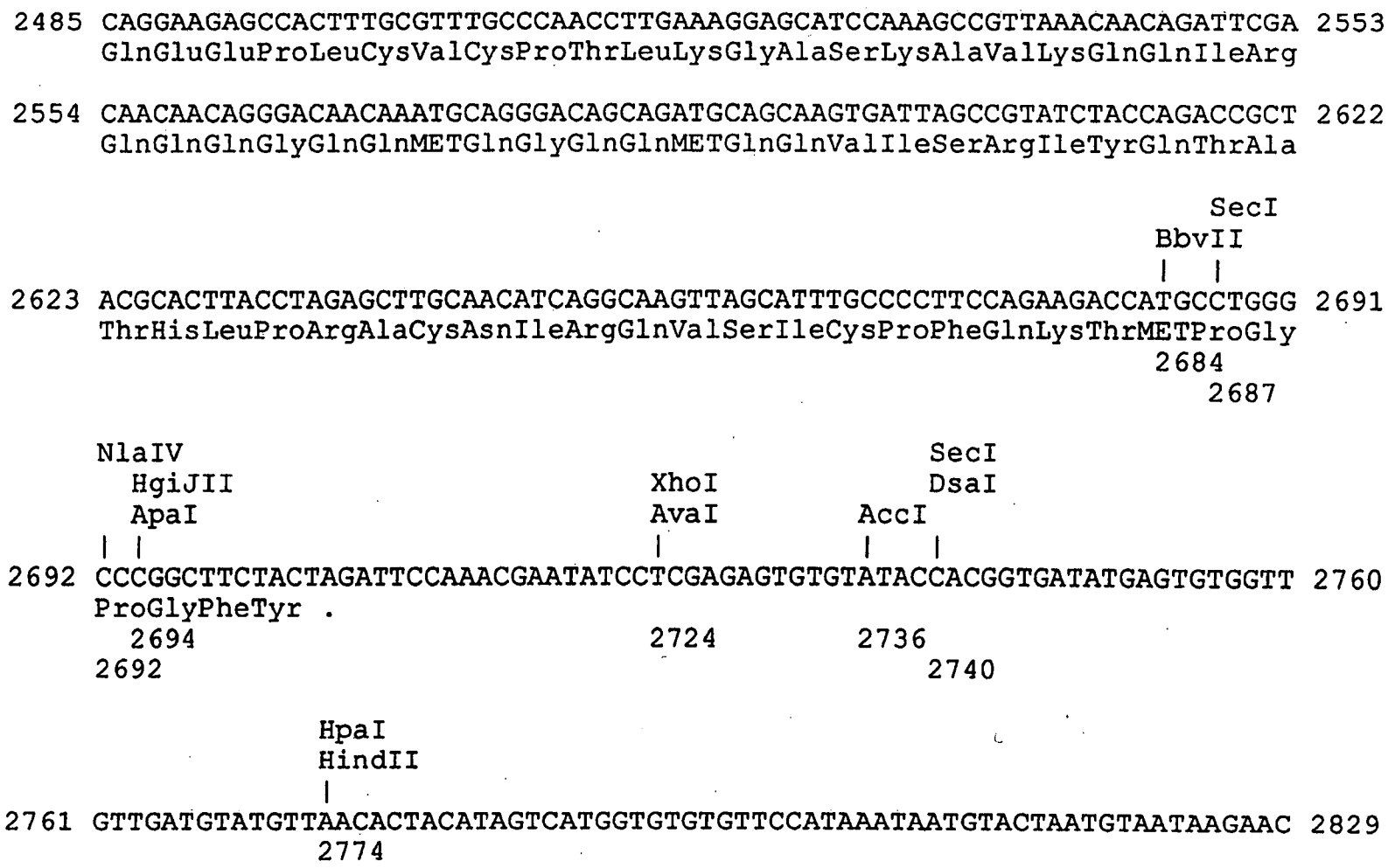



FIGURE 10
Page 8 of 13

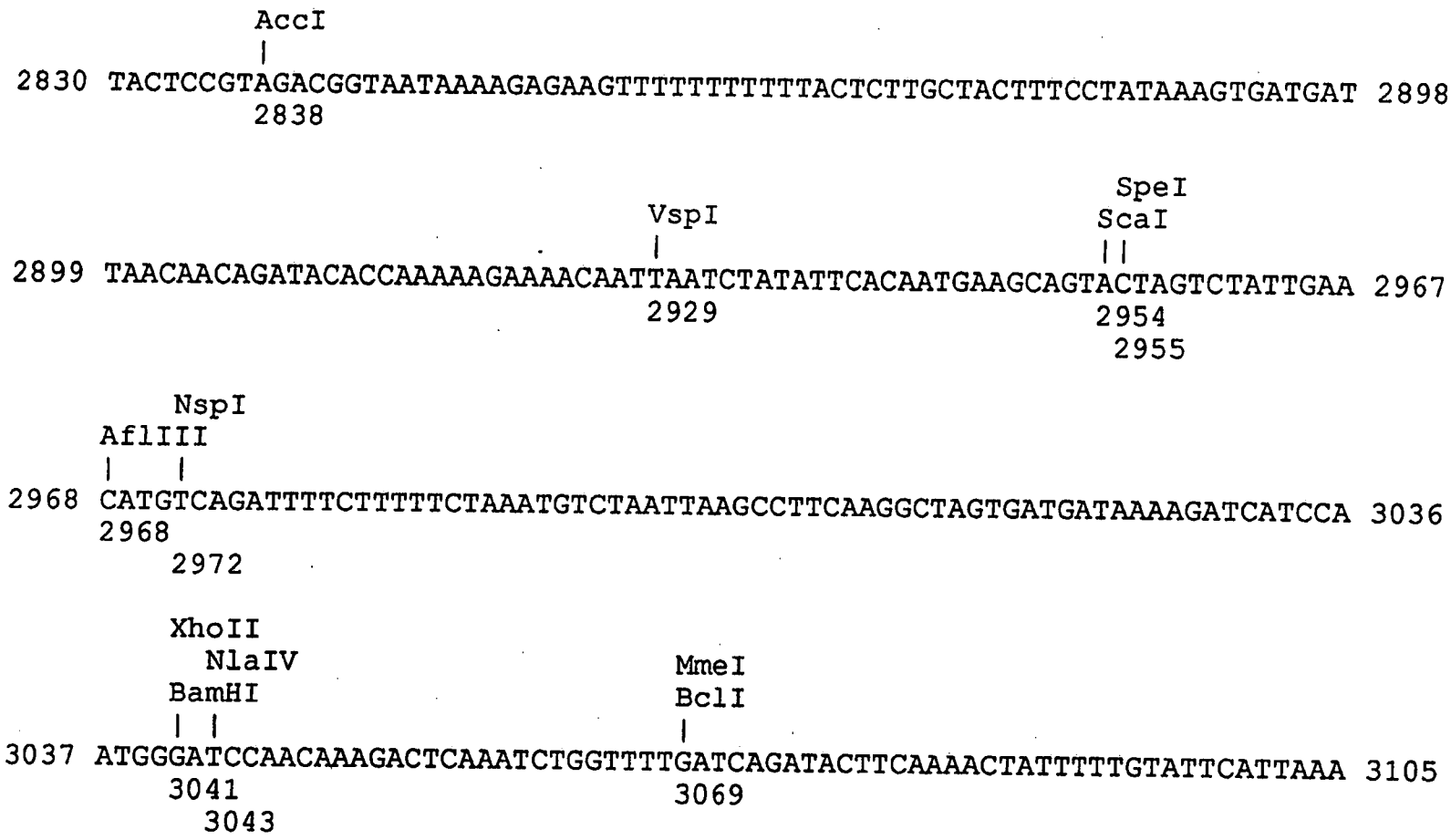



FIGURE 10
Page 9 of 13

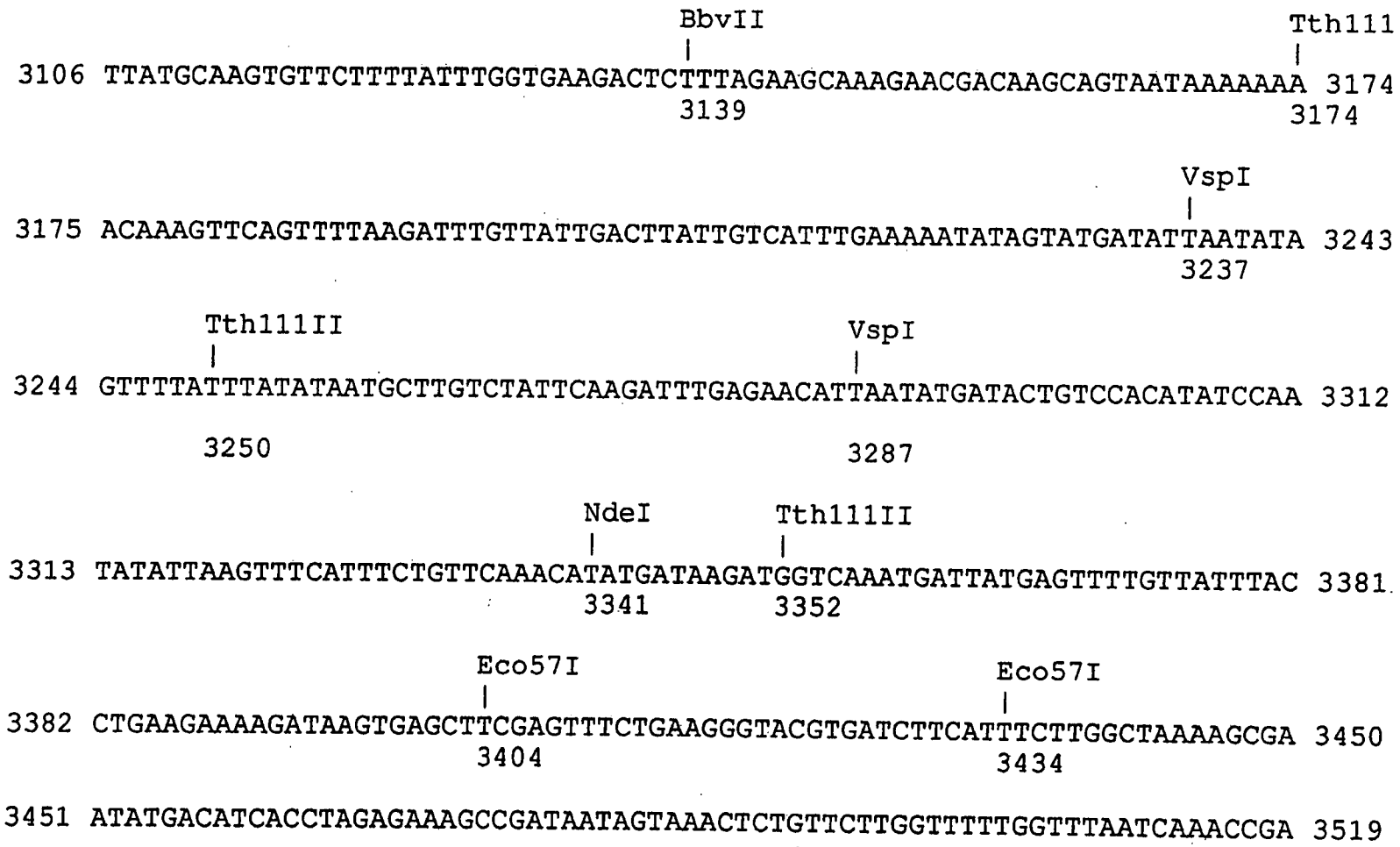



FIGURE 10
Page 10 of 13

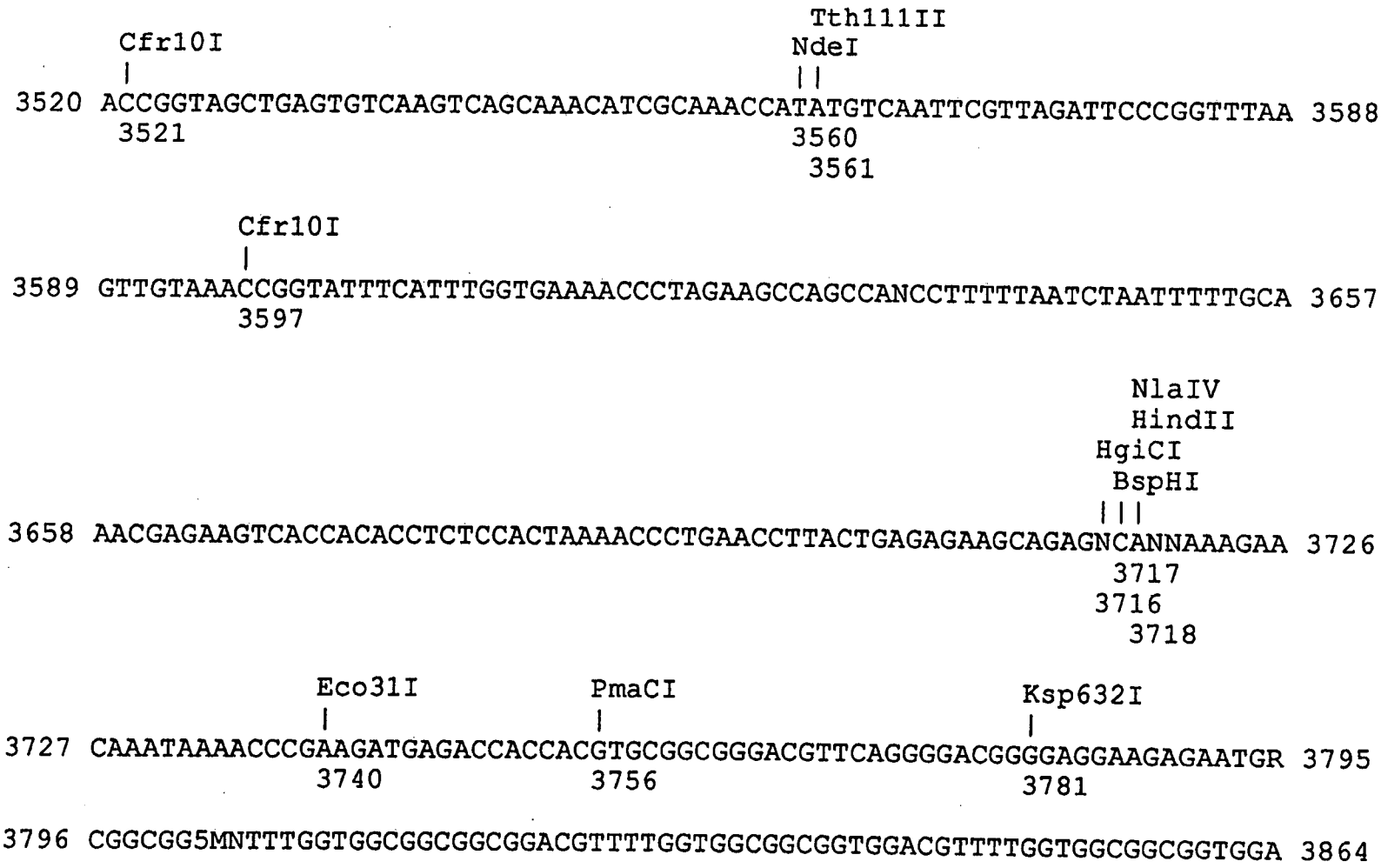



FIGURE 10
Page 11 of 13

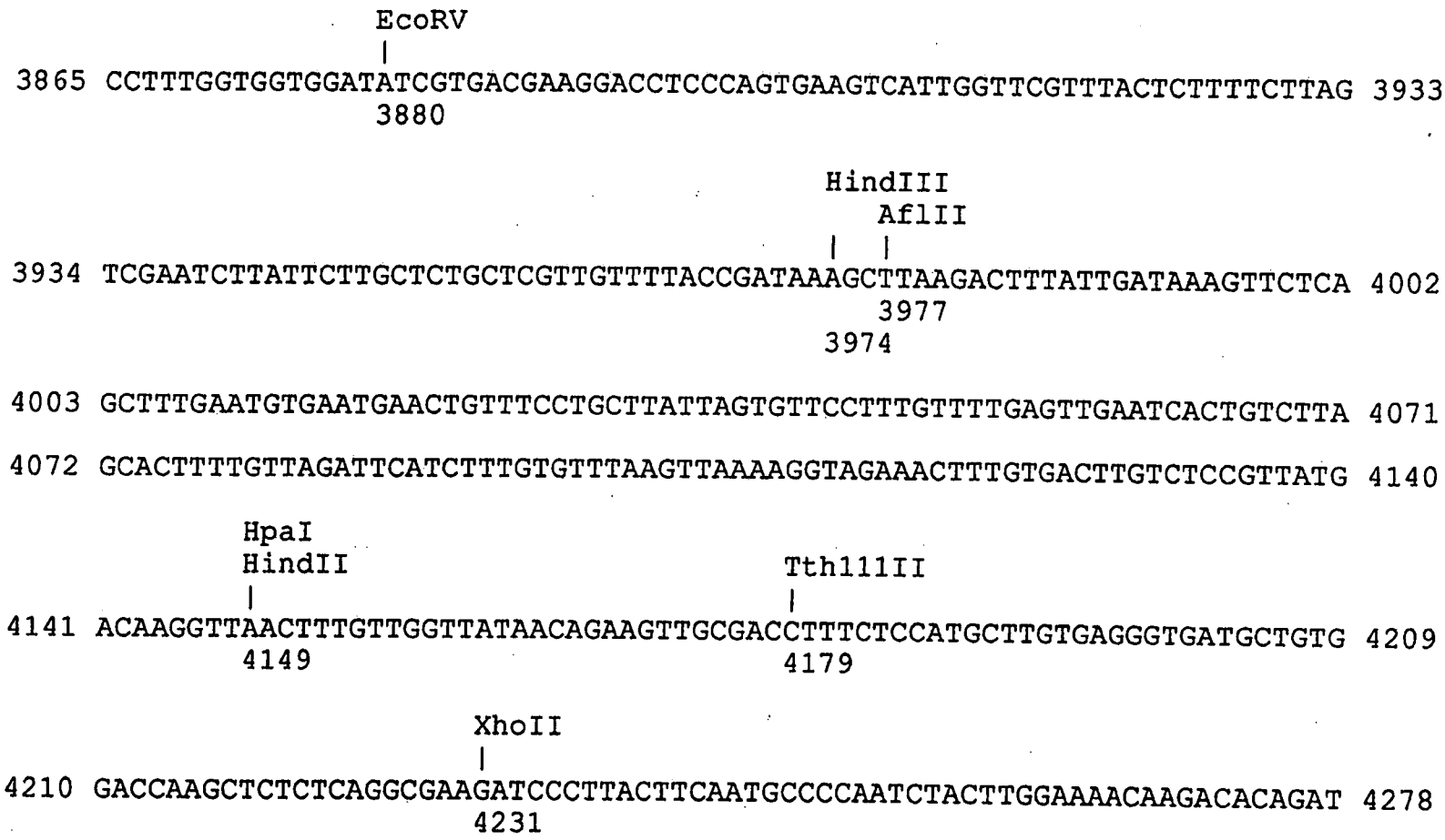

EcoRV
3865 CCTTTGGTGGTGGATATCGTGACGAAGGACCTCCCAGTGAAGTCATTGGTTCGTTTACTCTTTTCTTAG 3933
3880
HindIII
AflII
3934 TCGAATCTTATTCTTGCTCTGCTCGTTGTTTTACCGATAAAGCTTAAGACTTTATTGATAAAGTTCTCA 4002
3977
3974
4003 GCTTTGAATGTGAATGAACTGTTTCCTGCTTATTAGTGTTCCTTTGTTTTGAGTTGAATCACTGTCTTA 4071
4072 GCACTTTTGTTAGATTCATCTTTGTGTTTAAGTTAAAAGGTAGAAACTTTGTGACTTGTCTCCGTTATG 4140
HpaI
HindII
Tth111II
4141 ACAAGGTTAACTTTGTTGGTTATAACAGAAGTTGCGACCTTTCTCCATGCTTGTGAGGGTGATGCTGTG 4209
4149
4179
XhoII
4210 GACCAAGCTCTCTCAGGCGAAGATCCCTTACTTCAATGCCCCAATCTACTTGGAAAACAAGACACAGAT 4278
4231

FIGURE 10

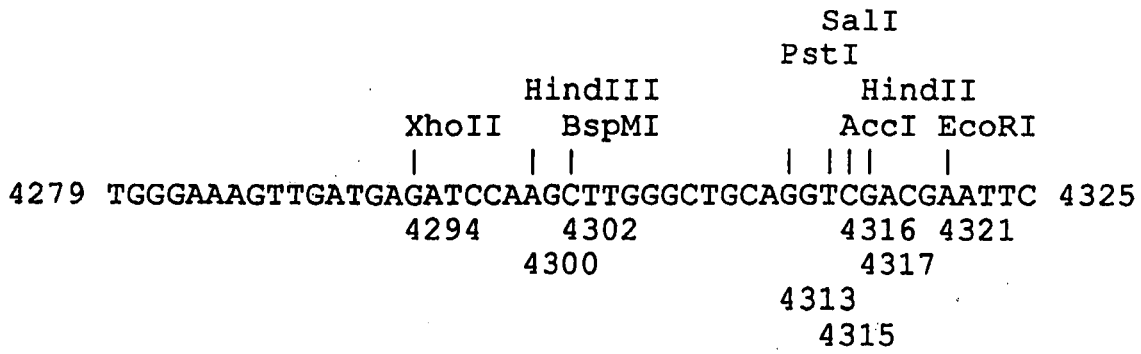



FIGURE 10
Page 13 of 13

# INTERNATIONAL SEARCH REPORT

International Application No. PCT/US91/01746

## I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(5): C12N 1/21, 15/29, 15/82; C07H 15/12

U.S. CL.: 435/172.3, 240.4, 252.3; 536/27

## II. FIELDS SEARCHED

Minimum Documentation Searched [7]

| Classification System | Classification Symbols |
|---|---|
| U.S. | 435/172.3, 240.4, 252.3 800/205, DIG.69 ; 536/27 |

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

USPTO AUTOMATED PATENT SYSTEM: DIALOG FILES BIOTECH AND PATENTS.
SEE ATTACHMENT FOR SEARCH TERMS

## III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

| Category [10] | Citation of Document, [11] with indication, where appropriate, of the relevant passages [12] | Relevant to Claim No. [13] |
|---|---|---|
| | SEE ATTACHED PAGES | |

[10] Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

## IV. CERTIFICATION

| Date of the Actual Completion of the International Search | Date of Mailing of this International Search Report |
|---|---|
| 24 June 1991 | 07 AUG 1991 |
| **International Searching Authority** | **Signature of Authorized Officer** |
| RO/US | P.Rhodes |

| Category * | Citation of Document, with indication, where appropriate, of the relevant passages | Relevant to Claim No |
|---|---|---|
| Y | World Soybean Research Conference III: Proceedings (Westview Press): Shibles (ed); Published 1985; Goodman et al; "Biotechnology and its impact on future improvements in soybean production and use"; pages 261-271. See pages 264-265. | 1-22, 34-37 |
| Y | Journal of Lipid Research; Volume 26; Issued 1985; Mattson et al; "Comparison of effects of dietary saturated, monounsaturated, and poly-unsaturated fatty acids on plasma and lipids and lipoprotein in man"; pages 194-202. See entire document. | 1-22, 34-37 |
| Y | EP A 0,255,377 (KRIDL et al) 03 February 1988. See entire document. | 1-22, 34-37 |
| Y | Trends in Biotechnology; Volume 5; Issued February 1987; Knauf; "The application of genetic engineering to oilseed crops"; pages 40-47. See entire document. | 1-22, 34-37 |
| Y | Trends in Biotechnology; Volume 7; Issued May 1989; Battey et al; "Genetic engineering for plant oils: potential and limitations"; pages 122-126. See entire document. | 1-22, 34-37 |
| Y | US, A, 4,446,235 (SEEBURG) 01 May 1984. See entire document. | 1-22, 34-37 |
| Y | US, A, 4,394,443 (WEISSMAN et al) 19 July 1983. See entire document. | 1-22, 34-37 |
| Y | Methods in Enzymology; Volume 71; Issued 1981; McKeon et al; "Stearoyl-acyl carrier protein desaturase from safflower seeds"; pages 275-281. See entire document. | 1-22, 34-37 |
| Y | Archives of Biochemistry and Biophysics; Volume 162; Issued 1974; Javorski et al; "Fat metabolism in higher plants, properties of a soluble stearyl-acyl carrier protein destaurase from maturing Carthamus tinctorius"; pages 158-165. See entire document. | 1-22, 34-37 |
| Y | The Journal of Biological Chemistry; Volume 257, Number 20; Issued 25 October 1982; McKeon et al; "Purification and characterization of the stearoyl-acyl carrier protein desaturase and the acyl-acyl carrier protein thioesterase from maturing seeds of safflower"; pages 12141-12147. See entire document. | 1-22, 34-37 |

| Category * | Citation of Document, with indication, where appropriate, of the relevant passages | Relevant to Claim No |
|---|---|---|
| X | Proceedings of the Flax Institute USA; Volume 41, Number 3; Issued 1971; Downey et al; "Genetic control of fatty acid composition in oilseed crops"; pages 1-3. See entire document. | 23-33, 38-45 |
| X<br>Y | EP, A, 0323753 (WONG et al) 12 July 1989. See entire document. | 23-29, 31-33<br>38-39, 41-45<br>30, 40 |
| X<br>Y | Journal of the American Oil Chemists Society; Volume 61, Number 1; Issued January 1984; Wilcox et al; "Genetic alteration of soybean oil composition by a chemical mutagen"; pages 97-100. See entire document. | 23, 25-27, 29,<br>31, 33, 38-45<br>30, 40 |
| X<br>Y | Journal of the American Oil Chemists Society; Volume 59, Number 5; Issued May 1982; Wolf et al; "Effect of temperature on soybean seed constituents: oil, protein, moisture, fatty acids, amino acids and sugars"; pages 230-232. See entire document. | 23, 25-27, 29,<br>31, 33, 38-45<br>30, 40 |
| Y | Lipids; Volume 4, Number 6; Issued 1969; Inkpen et al; "Desaturation of palmitate and stearate by cell-free fractions from soybean cotyledons"; pages 539-543. See entire document. | 30, 40 |
| Y | The Journal of Biological Chemistry; Volume 241; Issued 25 April 1966; Nagai et al; "Enzymatic desaturation of stearyl acyl carrier protein"; pages 1925-1927. See entire document. | 30, 40 |
| X<br>Y | The Journal of Heredity; Volume 80, Number 3; Issued March 1989; Moore et al; "The inheritance of high oleic acid in peanut"; pages 252-253. See entire document. | 23, 25-27, 29<br>31, 33, 38-45<br>30, 40 |
| X<br>Y | Crop Science; Volume 24; Issued November-December 1984; Carver et al; "Developmental changes in acyl-composition of soybean seed selected for high oleic acid concentration"; pages 1016-1019. See entire document. | 23, 25-27, 29<br>31, 33, 38-45<br>30, 40 |
| X<br>Y | Bodman et al., "Soybeans and Soybean Products: Processing of edible soybean oil" published 1951 by Interscience Publishers, Inc. (N.Y.), pages 649-725, see only pages 702-709. | 31 and 33<br>32 |

# EXHIBIT F

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 616 644 B1

# (12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
**02.07.2003 Bulletin 2003/27**

(21) Application number: **93900716.7**

(22) Date of filing: **03.12.1992**

(51) Int Cl.[7]: **C12N 15/53**, C12N 15/82, C11B 1/00, C12Q 1/68

(86) International application number:
**PCT/US92/10284**

(87) International publication number:
**WO 93/011245 (10.06.1993 Gazette 1993/14)**

## (54) FATTY ACID DESATURASE GENES FROM PLANTS

FETTSÄURE-DESATURASE GENE AUS PFLANZEN

GENES DE DESATURASE D'ACIDES GRAS A PARTIR DE PLANTES

(84) Designated Contracting States:
**DE DK FR GB IT NL SE**

(30) Priority: **04.12.1991 US 804259**

(43) Date of publication of application:
**28.09.1994 Bulletin 1994/39**

(73) Proprietor: **E.I. DU PONT DE NEMOURS AND COMPANY**
**Wilmington Delaware 19898 (US)**

(72) Inventors:
- **BROWSE, John**
  **Palouse, WA 99161 (US)**
- **GRAU, Luis, Perez**
  **Wilmington, DE 19809 (US)**
- **KINNEY, Anthony, J.**
  **Wilmington, DE 19809 (US)**
- **PIERCE, John, W., Jr.**
  **Wilmington, DE 19803 (US)**
- **WIERZBICKI, Anna, M.**
  **Wilmington, DE 19805 (US)**
- **YADAV, Narendra, S.**
  **Wilmington, DE 19810 (US)**

(74) Representative: **Grant, Anne Rosemary et al**
**Frank B. Dehn & Co.**
**179 Queen Victoria Street**
**London EC4V 4EL (GB)**

(56) References cited:
**WO-A-91/13972**

- **UCLA SYMP. MOL. CELL. BIOL.; NEW SER., PLANT GENE TRANSFER vol. 129, 1990, pages 301 - 309 BROWSE, J., ET AL. 'Strategies for modifying plant lipid composition'**
- **SCIENCE vol. 252, 5 April 1991, LANCASTER, PA US pages 80 - 87 SOMERVILLE, C., ET AL. 'Plant lipids: Metabolism, mutants, and membranes'**
- **THEOR. APPL. GENET. vol. 80, no. 2, 1990, pages 234 - 240 LEMIEUX, B., ET AL. 'Mutants of Arabidopsis with alterations in seed lipid fatty acid composition'**
- **SCIENCE vol. 258, 20 November 1992, LANCASTER, PA US pages 1353 - 1355 ARONDEL, V., ET AL. 'Map-based cloning of a gene controlling omega-3 fatty acid desaturation in Arabidopsis'**
- **PLANT PHYSIOLOGY. vol. 81, no. 3, 1986, ROCKVILLE, MD, USA. pages 859 - 864 BROWSE, J., ET AL. 'A mutant of Arabidopsis deficient in C18:3 and C16:3 leaf lipids'**



EP 0 616 644 B1

## Description

### FIELD OF THE INVENTION

**[0001]** The invention relates to the preparation and use of nucleic acid fragments encoding fatty acid desaturase enzymes to modify plant lipid composition.

### BACKGROUND OF THE INVENTION

**[0002]** Plant lipids have a variety of industrial and nutritional uses and are central to plant membrane function and climatic adaptation. These lipids represent a vast array of chemical structures, and these structures determine the physiological and industrial properties of the lipid. Many of these structures result either directly or indirectly from metabolic processes that alter the degree of unsaturation of the lipid. Different metabolic regimes in different plants produce these altered lipids, and either domestication of exotic plant species or modification of agronomically adapted species is usually required to economically produce large amounts of the desired lipid.

**[0003]** Plant lipids find their major use as edible oils in the form of triacylglycerols. The specific performance and health attributes of edible oils are determined largely by their fatty acid composition. Most vegetable oils derived from commercial plant varieties are composed primarily of palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2) and linolenic (18:3) acids. Palmitic and stearic acids are, respectively, 16- and 18-carbon-long, saturated fatty acids. Oleic, linoleic, and linolenic acids are 16-carbon-long, unsaturated fatty acids containing one, two, and three double bonds, respectively. Oleic acid is referred to as a mono-unsaturated fatty acid, while linoleic and linolenic acids are referred to as poly-unsaturated fatty acids. The relative amounts of saturated and unsaturated fatty acids in commonly used, edible vegetable oils are summarized below (Table 1):

TABLE 1

| Percentages of Saturated and Unsaturated Fatty Acids in the Oils of Selected Oil Crops | | | |
|---|---|---|---|
| | Saturated | Mono-unsaturated | Poly-unsaturated |
| Canolla | 6% | 58% | 36% |
| Soybean | 15% | 24% | 61% |
| Corn | 13% | 25% | 62% |
| Peanut | 18% | 48% | 34% |
| Safflower | 9% | 13% | 78% |
| Sunflower | 9% | 41% | 51% |
| Cotton | 30% | 19% | 51% |

**[0004]** Many recent research efforts have examined the role that saturated and unsaturated fatty acids play in reducing the risk of coronary heart disease. In the past, it was believed that mono-unsaturates, in contrast to saturates and poly-unsaturates, had no effect on serum cholesterol and coronary heart disease risk. Several recent human clinical studies suggest that diets high in mono-unsaturated fat and low in saturated fat may reduce the "bad" (low-density lipoprotein) cholesterol while maintaining the "good" (high-density lipoprotein) cholesterol (Mattson et al., Journal of Lipid Research (1985) 26:194-202).

**[0005]** A vegetable oil low in total saturates and high in mono-unsaturates would provide significant health benefits to consumers as well as economic benefits to oil processors. As an example, canola oil is considered a very healthy oil. However, in use, the high level of poly-unsaturated fatty acids in canola oil renders the oil unstable, easily oxidized, and susceptible to development of disagreeable odors and flavors (Gailliard, 1980, Vol. 4, pp. 85-116 In: Stumpf, P. K., Ed., The Biochemistry of Plants, Academic Press, New York). The levels of poly-unsaturates may be reduced by hydrogenation, but the expense of this process and the concomitant production of nutritionally questionable trans isomers of the remaining unsaturated fatty acids reduces the overall desirability of the hydrogenated oil (Mensink et al., New England J. Medicine (1990) N323: 439-445). Similar problems exist with soybean and corn oils.

**[0006]** For specialized uses, high levels of poly-unsaturates can be desirable. Linoleate and linolenate are essential fatty acids in human diets, and an edible oil high in these fatty acids can be used for nutritional supplements, for example in baby foods. Linseed oil, derived from the Flax plant (Linum usitatissimum), contains over 50% linolenic acid and has widespread use in domestic and industrial coatings since the double bonds of the fatty acids react rapidly with oxygen to polymerize into a soft and flexible film. Although the oil content of flax is comparable to canola (around 40% dry weight of seed), high yields are only obtained in warm temperatures or subtropical climates. In the USA flax is highly susceptible to rust infection. It will be commercially useful if a crop such as soybean or canola could be genetically

transformed by the appropriate desaturase gene(s) to synthesize oils with a high linolenic acid content.

**[0007]** Mutation-breeding programs have met with some success in altering the levels of poly-unsaturated fatty acid levels found in the edible oils of agronomic species. Examples of commercially grown varieties are high (85%) oleic sunflower and low (2%) linolenic flax (Knowles, (1980) pp. 35-38 In: Applewhite, T. H., Ed., World Conference on Biotechnology for the Fats and Oils Industry Proceedings, American Oil Chemists' Society). Similar commercial progress with the other plants shown in Table 1 has been largely elusive due to the difficult nature of the procedure and the pleiotropic effects of the mutational regime on plant hardiness and yield potential.

**[0008]** The biosynthesis of the major plant lipids has been the focus of much research (Browse et al., Ann. Rev. Plant Physiol. Mol. Biol. (1991) 42:467-506). These studies show that, with the notable exception of the soluble stearoyl-acyl carrier protein desaturase, the controlling steps in the production of unsaturated fatty acids are largely catalyzed by membrane-associated fatty acid desaturases. Desaturation reactions occur in plastids and in the endoplasmic reticulum using a variety of substrates including galactolipids, sulfolipids, and phospholipids. Genetic and physiological analyses of Arabidopsis thaliana nuclear mutants defective in various fatty acid desaturation reactions indicates that most of these reactions are catalyzed by enzymes encoded at single genetic loci in the plant. The analyses show further that the different defects in fatty acid desaturation can have profound and different effects on the ultra-structural morphology, cold sensitivity, and photosynthetic capacity of the plants (Ohlrogge, et al., Biochim. Biophys. Acta (1991) 1082:1-26). However, biochemical characterization of the desaturase reactions has been meager. The instability of the enzymes and the intractability of their proper assay has largely limited researchers to investigations of enzyme activities in crude membrane preparations. These investigations have, however, demonstrated the role of delta-12 desaturase and delta-15 desaturase activities in the production of linoleate and linolenate from 2-oleoyl-phosphatidylcholine and 2-linoleoyl-phosphatidylcholine, respectively (Wang et al., Plant Physiol. Biochem. (1988) 26:777-792). Thus, modification of the activities of these enzymes represents an attractive target for altering the levels of lipid unsaturation by genetic engineering.

**[0009]** Genes from plants for stearoyl-acyl carrier protein desaturase, the only soluble fatty acid desaturase known, have been described (Thompson, et al., Proc. Natl. Acad. Sci. U.S.A. (1991) 88:2578-2582; Shanklin et al., Proc. Natl. Acad. Sci. USA (1991) 88:2510-2514). Stearoyl-coenzyme-A desaturase genes from yeast, rat, and mice have also been described (Stukey, et al., J. Biol. Chem.(1990) 265:20144-20149; Thiede, et al., J. Biol. Chem. (1986) 261: 13230-13235; Kaestner, et al., J. Biol. Chem. (1989) 264:14755-1476). No evidence exists in the public art that describes the isolation of fatty acid desaturases other than stearoyl-ACP desaturases from higher plants or their corresponding genes. A fatty acid desaturase gene from the cyanobacterium, Synechocystis PCC 6803, has also been described (Wada, et al., Nature (1990) 347:200-203). This gene encodes a fatty acid desaturase, designated des A, that catalyzes the conversion of oleic acid at the 1 position of galactolipids to linoleic acid. However, these genes have not proven useful for isolating plant fatty acid desaturases other than stearoyl-ACP desaturase via sequence-dependent protocols, and the present art does not indicate how to obtain plant fatty acid desaturases other than stearoyl-ACP desaturases or how to obtain fatty acid desaturase-related enzymes. Thus, the present art does not teach how to obtain glycerolipid desaturases from plants. Furthermore, there is no evidence that a method to control the nature and levels of unsaturated fatty acids in plants using nucleic acids encoding fatty acid desaturases other than stearoyl-ACP desaturase is known in the art.

**[0010]** The biosynthesis of the minor plant lipids has been less well studied. While hundreds of different fatty acids have been found, many from the plant kingdom, only a tiny fraction of all plants have been surveyed for their lipid content (Gunstone, et al., Eds., (1986) The Lipids Handbook, Chapman and Hall Ltd., Cambridge). Accordingly, little is known about the biosynthesis of these unusual fatty acids and fatty acid derivatives. Interesting chemical features found in such fatty acids include, for example, allenic and conjugated double bonds, acetylenic bonds, trans double bonds, multiple double bonds, and single double bonds in a wide number of positions and configurations along the fatty acid chain. Similarly, many of the structural modifications found in unusual lipids (e.g., hydroxylation, epoxidation, cyclization, etc.) are probably produced via further metabolism following chemical activation of the fatty acid by desaturation or they involve a chemical reaction that is mechanistically similar to desaturation. For example, evidence for the mechanism of hydroxylation of fatty acids being part of a general mechanism of enzyme-catalyzed desaturation in eukaryotes has been obtained by substituting a sulfur atom in the place of carbon at the delta-9 position of stearic acid. When incubated with yeast cell extracts the thiostearate was converted to a 9-sulfoxide (Buist et al. (1987) Tetrahedron Letters 28:857-860). This sulfoxidation was specific for sulfur at the delta-9 position and did not occur in a yeast delta-9-desaturase deficient mutant (Buist & Marecak (1991) Tetrahedron Letters 32:891-894). The 9-sulfoxide is the sulfur analogue of 9-hydroxyoctadecastearate, the proposed intermediate of stearate desaturation. Thus fatty-acid desaturase cDNAs may serve as useful probes for cDNAs encoding fatty-acid hydroxylases and other cDNAs which encode enzymes with reaction mechanisms similar to fatty-acid desaturation. Many of these fatty acids and derivatives having such features within their structure could prove commercially useful if an agronomically viable species could be induced to synthesize them by introduction of a gene encoding the appropriate desaturase.

SUMMARY OF THE INVENTION

**[0011]** Applicants have discovered a means to control the nature and levels of unsaturated fatty acids in plants. Nucleic acid fragments from glycerolipid desaturase cDNAs or genes are used to create chimeric genes. The chimeric genes may be used to transform various plants to modify the fatty acid composition of the plant or the oil produced by the plant. One embodiment of the invention is an isolated nucleic acid fragment comprising a nucleic acid sequence encoding a fatty acid desaturase which hybridises to the nucleotide sequence set forth in SEQ ID NO: 1 under one of the following sets of conditions: (a) hybridisation in 50 mM Tris-HCl, pH 7.5, 1M NaCl, 1% sodium dodecyl sulfate (SDS), 5% dextran sulfate and 0.1 mg/mL denatured salmon sperm DNA at 50°C and wash twice at room temperature with 2X SSPE 0.1% SDS for 5 minutes and 10 minutes, followed by washing for 5 minutes at 50°C in 0.5X SSPE 0.1% SDS; (b)hybridisation in 50 mM Tris-HCl, pH 7.5, 1M NaCl, 1% sodium dodecyl sulfate (SDS), 5% dextran sulfate and 0.1 mg/mL denatured salmon sperm DNA at 50°C and wash twice at room temperature with 2X SSPE, 1% SDS for 5 minutes, then washing for 5 minutes at 50°C in 0.2X SSPE, 1% SDS; or (c) hybridisation in 50 mM Tris, pH 7.6, 6X SSC, 5X Denhardt's, 0.5% sodium dodecyl sulfate (SDS), 100 µg denatured calf thymus DNA at 50°C and wash with 6X SSC, 0.5% SDS at room temperature for 15 minutes, repeat with 2X SSC, 0.5% SDS at 45°C for 30 minutes, then repeat twice with 0.2X SSC, 0.5% SDS at 50°C for 30 minutes each. More specifically, a preferred embodiment of the invention is an isolated nucleic acid fragment comprising a nucleotide sequence encoding a plant delta-15 fatty acid desaturase or a fatty acid desaturase-related enzyme with an amino acid identity of 50%, 65%, 90% or greater to the polypeptide encoded by SEQ ID NOS: 1, 4, 6, 8, 10, 12, 14, or 16. The isolated fragment in these embodiments is isolated from a plant selected from the group consisting of soybean, oilseed Brassica species, Arabidopsis thaliana and corn.

**[0012]** Another embodiment of this invention involves the use of these nucleic acid fragments in sequence-dependent protocols. Examples include use of the fragments as hybridization probes to isolate other glycerolipid desaturase cDNAs or genes. A related embodiment involves using the disclosed sequences for amplification of DNA fragments encoding other glycerolipid desaturases.

**[0013]** Another aspect of this invention involves chimeric genes capable of causing altered levels of the linolenic acid in a transformed plant cell, the gene comprising nucleic acid fragments encoding a plant delta-15 fatty acid desaturase or a fatty acid desaturase-related enzyme of the invention, preferably with an amino acid identity of 50%, 65%, 90% or greater to the polypeptide encoded by SEQ ID NOS:1, 4, 6, 8, 10, 12, 14, or 16 operably linked in suitable orientation to suitable regulatory sequences. Preferred are those chimeric genes which incorporate nucleic acid fragments encoding delta-15 fatty acid desaturase cDNAs or genes. Plants from seeds of plants containing the chimeric genes described are also claimed.

**[0014]** Yet another embodiment of the invention involves a method of producing seed oil containing altered levels of linolenic (18:3) acid comprising: (a) transforming a plant cell with a chimeric gene described above; (b) growing fertile plants from the transformed plant cells of step (a); (c) screening progeny seeds from the fertile plants of step (b) for the desired levels of linolenic (18:3) acid, and (d) processing the progeny seed of step (c) to obtain seed oil containing altered levels of the unsaturated fatty acids. Preferred plant cells and oils are derived from soybean, rapeseed, sunflower, cotton, cocoa, peanut, safflower, coconut, flax, oil palm, and corn. Preferred methods of transforming such plant cells would include the use of Ti and Ri plasmids of Agrobacterium, electroporation, and high-velocity ballistic bombardment.

**[0015]** The invention may also be used in a method of breeding plant species to obtain altered levels of poly-unsaturated fatty acids, specifically linolenic (18:3) acid in seed oil of oil-producing plants. This method involves (a) making a cross between two varieties of an oilseed plant differing in the linolenic acid trait; (b) making a Southern blot of restriction enzyme digested genomic DNA isolated from several progeny plants resulting from the cross of step (a); and (c) hybridizing the Southern blot with the radiolabeled nucleic acid fragments encoding the claimed glycerolipid desaturases.

**[0016]** The invention is also embodied in a method of RFLP mapping that uses the isolated Arabidopsis thaliana delta-15 desaturase sequences described herein.

**[0017]** The invention is also embodied in plants capable of producing altered levels of glycerolipid desaturase by virtue of containing the chimeric genes described herein. Further, the invention may be used to provide seed oil obtained from such plants.

**[0018]** The invention is also embodied in a method of RFLP mapping in a genomic RFLP marker comprising(a) making a cross between two varieties of plants; (b)making a Southern blot of restriction enzyme digested genomic DNA isolated from several progeny plants resulting from the cross of step (a); and (c) hybridizing the Southern blot with a radiolabelled nucleic acid fragments of the claimed fragments.

**[0019]** The invention may also be used in a method to isolate nucleic acid fragments encoding fatty acid desaturases and fatty acid desaturase-related enzymes, comprising (a)comparing SEQ ID NOS:2, 5, 7, 9, 11, 13, 15 and 17 with other fatty acid desaturase polypeptide sequences; (b) identifying the conserved sequence(s) of 4 or more amino acids

obtained in step a; (c) making region-specific nucleotide probe(s) or oligomer(s) based on the conserved sequences identified in step b; and d) using the nucleotide probe(s) or oligomers(s) of step c to isolate sequences encoding fatty acid desaturases and fatty-acid desaturase-related enzymes by sequence-dependent protocols.

## BRIEF DESCRIPTION OF THE SEQUENCE DESCRIPTIONS

**[0020]** The invention can be more fully understood from the following detailed description and the Sequence Descriptions which form a part of this application. The Sequence Descriptions contain the one letter code for nucleotide sequence characters and the three letter code for amino acids in conformity with the IUPAC-IUB standard described in Nucleic Acids Research 13:3021-3030 (19085) and 37 C.F.R. 1.822 which are incorporated herein by reference.

**[0021]** SEQ ID NO:1 shows the complete 5' to 3' nucleotide sequence of 1350 base pairs of the Arabidopsis cDNA which encodes delta-15 desaturase in plasmid pCF3. Nucleotides 46 to 48 are the putative initiation codon of the open reading frame (nucleotides 46 to 1206). Nucleotides 1204 to 1206 are the termination codon. Nucleotides 1 to 45 and 1207 to 1350 are the 5' and 3' untranslated nucleotides, respectively. The 386 amino acid protein sequence in SEQ ID NO:1 is that deduced from the open reading frame.

**[0022]** SEQ ID NO:2 is the deduced peptide of the open-reading frame of SEQ ID NO:1.

**[0023]** SEQ ID NO:3 is a partial nucleotide sequence of the Arabidopsis genomic DNA insert in plasmid pF1 which shows the genomic sequence in the region of the Arabidopsis genome that encodes delta-15 desaturase. Nucleotides 68-255 are identical to nucleotides 1-188 of SEQ ID NO:1. Nucleotides 47 to 49 and 56 to 58 are termination codons in the same reading frame as the open reading frame in SEQ ID NO:1.

**[0024]** SEQ ID NO:4 shows the 5' to 3' nucleotide sequence of the insert in plasmid pACF2-2 of 1525 base pairs of the Arabidopsis thaliana cDNA that encodes a plastid delta-15 fatty acid desaturase. Nucleotides 10-12 and nucleotides 1348 to 1350 are, respectively, the putative initiation codon and the termination codon of the open reading frame (nucleotides 10 to 1350). Nucleotides 1 to 9 and 1351 to 1525 are, respectively, the 5' and 3' untranslated nucleotides.

**[0025]** SEQ ID NO:5 is the deduced peptide of the open reading frame of SEQ ID NO:4.

**[0026]** SEQ ID NO:6 shows the complete 5' to 3' nucleotide sequence of 1336 base pairs of the Brassica napus seed cDNA, found in plasmid pBNSF3-2, which encodes a microsomal delta-15 glycerolipid desaturase. Nucleotides 79 to 82 are the putative initiation codon of the open reading frame (nucleotides 79 to 1212). Nucleotides 1210 to 1212 are the termination codon. Nucleotides 1 to 78 and 1213 to 1336 are the 5' and 3' unstranslated nucleotides respectively.

**[0027]** SEQ ID NO:7 is the deduced peptide of the open reading frame of SEQ ID NO:6.

**[0028]** SEQ ID NO:8 is the complete 5' to 3' nucleotide sequence of 1416 base pairs of the Brassica napus seed cDNA found in plasmid pBNSFd-2 which encodes a plastid delta-15 glycerolipid desaturase. Nucleotides 1 to 1215 correspond to a continuous open reading frame of 404 amino acids. Nucleotides 1213 to 1215 are the termination codon. Nucleotides 1215 to 1416 are the 3' untranslated nucleotides.

**[0029]** SEQ ID NO:9 is the deduced peptide of the open reading frame of SEQ ID NO:8.

**[0030]** SEQ ID NO:10 is the complete nucleotide sequence of the soybean (glycine max) microsomal delta-15 desaturase cDNA, found in plasmid pXF1, which the 2184 nucleotides of this sequence contain both the coding sequence and the 5' and 3' non-translated regions of the cDNA. Nucleotides 855 to 857 are the putative initiation codon of the open reading frame (nucleotides 855 to 2000). Nucleotides 1995 to 1997 are the termination codon. Nucleotides 1 to 854 and 1998 to 2184 are the 5' and 3' unstranslated nucleotides respectively. The 380 amino acid protein sequence in SEQ ID NO:7 is that deduced from the open reading frame.

**[0031]** SEQ ID NO:11 is the deduced peptide of the open reading frame in SEQ ID NO:10.

**[0032]** SEQ ID NO:12 is the complete 5' to 3' nucleotide sequence of 1676 base pairs of the soybean (Glycine max) seed cDNA found in plasmid pSFD-118bwp which encodes a soybean plastid delta-15 desaturase. Nucleotides 169 to 1530 correspond to a continuous open reading frame of 453 amino acids. Nucleotides 169 to 171 are the putative initiation codon of the open reading frame. Nucleotides 1528 to 1530 are the termination codon. Nucleotides 1531 to 1676 are the 3' untranslated nucleotides. Nucleotides 169 to 382 encode the putative plastid transit peptide, based on comparison of the deduced peptide with the soybean microsomal delta-15 peptide.

**[0033]** SEQ ID NO:13 is the deduced peptide of the open reading frame in SEQ ID NO:12.

**[0034]** SEQ ID NO:14 is the complete nucleotide sequence of a 396 bp polymerase chain reaction product derived from corn seed mRNA that is found in the insert of plasmid pPCR20. Nucleotides 1 to 31 and 364 to 396 correspond to the amplification primers described in SEQ ID NO:18 and SEQ ID NO:19, respectively. Nucleotides 31 to 363 encode an internal region of a corn seed delta-15 desaturase that is 61.9% identical to the region between amino acids 137 and 249 of the Brassica napus delta-15 desaturase peptide sequence shown in SEQ ID NO:7.

**[0035]** SEQ ID NO:15 is the deduced amino acid sequence of SEQ ID NO:14.

**[0036]** SEQ ID NO:16 shows the partial composite 5' to 3' nucleotide sequence of 472 bp derived from the inserts in plasmids pFadx-2 and pYacp7 for Arabidopsis thaliana cDNA that encodes a plastid delta-15 fatty acid desaturase. Nucleotides 2-4 and nucleotides 468 to 470 are, respectively, the first and the last codons in the open reading frame.

[0037] SEQ ID NO:17 is deduced partial peptide sequence of the open reading frame in SEQ ID NO:16.

[0038] SEQ ID NO:18 One hundred and twenty eight fold degenerate sense 31-mer PCR primer. Nucleotides 1 to 8 correspond to the Bam H1 restriction enzyme recognition sequence. Nucleotides 9 to 137 correspond to amino acid residues 130 to 137 of SEQ ID NO:6 with a deoxyinosine base at nucleotide 11.

[0039] SEQ ID NO:19 Two thousand and forty eight-fold degenerate antisense 35-mer PCR primer. Nucleotides 1 to 8 correspond to the Bam H1 restriction enzyme recognition sequence. Nucleotides 9 to 35 correspond to amino acid residues 249 to 256 of SEQ ID NO:6 with a deoxyinosine base at nucleotide 15.

[0040] SEQ ID NO:20 Sixteen-fold degenerate sense 36-mers made to amino acid residues 97-108 in SEQ ID NO:2.

[0041] SEQ ID NO:21 Sixteen-fold degenerate sense 36-mers made to amino acid residues 97-108 in SEQ ID NO:2.

[0042] SEQ ID NO:22 Seventy two-fold degenerate sense 18-mers made to amino acid residues 100-105 in SEQ ID NO:2.

[0043] SEQ ID NO:23 Seventy two-fold degenerate sense 18-mers made to amino acid residues 100-105 in SEQ ID NO:2.

[0044] SEQ ID NO:24 Seventy two-fold degenerate antisense 18-mers made to amino acid residues 299-304 in SEQ ID NO:2.

[0045] SEQ ID NO:25 Seventy two-fold degenerate antisense 18-mers made to amino acid residues 299-304 in SEQ ID NO:2.

[0046] SEQ ID NO:26 Seventy two-fold degenerate antisense 18-mers made to amino acid residues 304-309 in SEQ ID NO:2.

[0047] SEQ ID NO:27 Seventy two-fold degenerate antisense 18-mers made to amino acid residues 304-309 in SEQ ID NO:2.

[0048] SEQ ID NO:28 Sixteen-fold degenerate sense 36-mers made to amino acid residues 97-108 in SEQ ID NO:2.

[0049] SEQ ID NO:29 Sixteen-fold degenerate sense 36-mers made to amino acid residues 97-108 in SEQ ID NO:2.

[0050] SEQ ID NO:30 Sixty four-fold degenerate antisense 38-mers made to amino acid residues 299-311 in SEQ ID NO:2.

[0051] SEQ ID NO:31 Sixty four-fold degenerate antisense 38-mers made to amino acid residues 299-311 in SEQ ID NO:2.

[0052] SEQ ID NO:32 A 135-mer made as an antisense strand to amino acid residues 97-141 in SEQ ID NO:2.

## DETAILED DESCRIPTION OF THE INVENTION

[0053] Applicants have isolated nucleic acid fragments that encode plant fatty acid desaturases and that are useful in modifying fatty acid composition in oil-producing species by transformation.

[0054] Thus, transfer of the nucleic acid fragments of the invention or a part thereof that encodes a functional enzyme, along with suitable regulatory sequences that direct the transciption of their mRNA, into a living cell will result in the production or over-production of plant fatty acid desaturases and will result in increased levels of unsaturated fatty acids in cellular lipids, including triacylglycerols.

[0055] Transfer of the nucleic acid fragments of the invention or a part thereof, along with suitable regulatory sequences that direct the transcription of their antisense RNA, into plants will result in the inhibition of expression of the endogenous fatty acid desaturase that is substantially homologous with the transferred nucleic acid fragment and will result in decreased levels of unsaturated fatty acids in cellular lipids, including triacylglycerols.

[0056] Transfer of the nucleic acid fragments of the invention or a part thereof, along with suitable regulatory sequences that direct the transciption of their mRNA, into plants may result in inhibition by cosuppression of the expression of the endogenous fatty acid desaturase gene that is substantially homologous with the transferred nucleic acid fragment and may result in decreased levels of unsaturated fatty acids in cellular lipids, including triacylglycerols.

[0057] The nucleic acid fragments of the invention can also be used as restriction fragment length polymorphism (RFLP) markers in Arabidopsis genetic mapping and plant breeding programs.

[0058] The nucleic acid fragments of the invention or oligomers derived therefrom can also be used to isolate other related glycerolipid desaturase genes using DNA, RNA, or a library of cloned nucleotide sequences from the same or different species by well known sequence-dependent protocols, including, for example, methods of nucleic acid hybridization and amplification by the polymerase chain reaction.

### Definitions

[0059] In the context of this disclosure, a number of terms shall be used. The term "fatty acid desaturase" used herein refers to an enzyme which catalyzes the breakage of a carbon-hydrogen bond and the introduction of a carbon-carbon double bond into a fatty acid molecule. The fatty acid may be free or esterified to another molecule including, but not limited to, acyl-carrier protein, coenzyme A, sterols and the glycerol moiety of glycerolipids. The term "glycerolipid

desaturases" used herein refers to a subset of the fatty acid desaturases that act on fatty acyl moieties esterified to a glycerol backbone. "Delta-12 desaturase" refers to a fatty acid desaturase that catalyzes the formation of a double bond between carbon positions 6 and 7 (numbered from the methyl end), (i.e., those that correspond to carbon positions 12 and 13 (numbered from the carbonyl carbon) of an 18 carbon-long fatty acyl chain or carbon positions 10 and 11 (numbered from the carbonyl carbon) of a 16 carbon-long fatty acyl chain). "Delta-15 desaturase" refers to a fatty acid desaturase that catalyzes the formation of a double bond between carbon positions 3 and 4 (numbered from the methyl end), (i.e., those that correspond to carbon positions 15 and 16 (numbered from the carbonyl carbon) of an 18 carbon-long fatty acyl chain and carbon positions 13 and 14 (numbered from the carbonyl carbon) of a 16 carbon-long fatty acyl chain). Examples of fatty acid desaturases include, but are not limited to, the microsomal delta-12 and delta-15 desaturases that act on phosphatidylcholine lipid substrates; the chloroplastic delta-12 and delta-15 desaturases that act on phosphatidyl glycerol and galactolipids; and other desaturases that act on such fatty acid substrates such as phospholipids, galactolipids, and sulfolipids. "Microsomal desaturase" refers to the cytoplasmic location of the enzyme, while "chloroplast desaturase" and "plastid desaturase" refer to the plastid location of the enzyme. These fatty acid desaturases may be found in a variety of organisms including, but not limited to, higher plants, diatoms, and various eukaryotic and prokaryotic microorganisms such as fungi and photosynthetic bacteria and algae. The term "homologous fatty acid desaturases" refers to fatty acid desaturases that catalyze the same desaturation on the same lipid substrate. Thus, microsomal delta-15 desaturases, even from different plant species, are homologous fatty acid desaturases. The term "heterologous fatty acid desaturases" refers to fatty acid desaturases that catalyze desaturations at different positions and/or on different lipid substrates. Thus, for example, microsomal delta-12 and delta-15 desaturases, which act on phosphatidylcholine lipids, are heterologous fatty acid desaturases, even when from the same plant. Similarly, microsomal delta-15 desaturase, which acts on phosphatidylcholine lipids, and chloroplast delta-15 desaturase, which acts on galactolipids, are heterologous fatty acid desaturases, even when from the same plant. It should be noted that these fatty acid desaturases have never been isolated and characterized as proteins. Accordingly the terms such as "delta-12 desaturase" and "delta-15 desaturase" are used as a convenience to describe the proteins encoded by nucleic acid fragments that have been isolated based on the phenotypic effects caused by their disruption. The term "fatty acid desaturase-related enzyme" refers to enzymes whose catalytic product may not be a carbon-carbon double bond but whose mechanism of action is similar to that of a fatty acid desaturase (that is, catalysis of the displacement of a carbon-hydrogen bond of a fatty acid chain to form a fatty-hydroxyacyl intermediate or end-product). This term is different from "related fatty acid desaturases", which refers to structural similarities between fatty acid desaturases.

**[0060]** The term "nucleic acid" refers to a large molecule which can be single-stranded or double-stranded, composed of monomers (nucleotides) containing a sugar, a phosphate and either a purine or pyrimidine. A "nucleic acid fragment" is a fraction of a given nucleic acid molecule. In higher plants, deoxyribonucleic acid (DNA) is the genetic material while ribonucleic acid (RNA) is involved in the transfer of the information in DNA into proteins. A "genome" is the entire body of genetic material contained in each cell of an organism. The term "nucleotide sequence" refers to the sequence of DNA or RNA polymers, which can be single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases capable of incorporation into DNA or RNA polymers. The term "oligomer" refers to short nucleotide sequences, usually up to 150 bases long. "Region-specific nucleotide probes" refers to isolated nucleic acid fragments derived from a cDNA or gene using a knowledge of the amino acid regions conserved between different fatty-acid desaturases which may be used to isolate cDNAS or genes for other fatty-acid desaturases or fatty acid desaturase-related enzymes using sequence dependent protocols. As used herein, the term "homologous to" refers to the relatedness between the nucleotide sequence of two nucleic acid molecules or between the amino acid sequences of two protein molecules. Estimates of such homology are provided by either DNA-DNA or DNA-RNA hybridization under conditions of stringency as is well understood by those skilled in the art (Hames and Higgins, Eds. (1985) Nucleic Acid Hybridisation, IRL Press, Oxford, U.K.); or by the comparison of sequence similarity between two nucleic acids or proteins, such as by the method of Needleman et al. (J. Mol. Biol. (1970) 48:443-453). As used herein, "substantially homologous" refers to nucleotide sequences that have more than 90% overall identity at the nucleotide level with the coding region of the claimed sequence, such as genes and pseudo-genes corresponding to the coding regions. The nucleic acid fragments described herein include molecules which comprise possible variations, both man-made and natural, such as but not limited to (a) those that involve base changes that do not cause a change in an encoded amino acid, or (b) which involve base changes that alter an amino acid but do not affect the functional properties of the protein encoded by the DNA sequence, (c) those derived from deletions, rearrangements, amplifications, random or controlled mutagenesis of the nucleic acid fragment, and (d) even occasional nucleotide sequencing errors.

**[0061]** "Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding) and following (3' non-coding) the coding region. "Fatty acid desaturase gene" refers to a nucleic acid fragment that expresses a protein with fatty acid desaturase activity. "Native" gene refers to an isolated gene with its own regulatory sequences as found in nature. "Chimeric gene" refers to a gene that comprises heterogeneous regulatory and coding sequences not found in nature. "Endogenous" gene refers to the native gene normally

found in its natural location in the genome and is not isolated. A "foreign" gene refers to a gene not normally found in the host organism but that is instead introduced by gene transfer. "Pseudo-gene" refers to a genomic nucleotide sequence that does not encode a functional enzyme.

**[0062]** "Coding sequence" refers to a DNA sequence that codes for a specific protein and excludes the non-coding sequences. It may constitute an "uninterrupted coding sequence", i.e., lacking an intron or it may include one or more introns bounded by appropriate splice junctions. An "intron" is a nucleotide sequence that is transcribed in the primary transcript but that is removed through cleavage and re-ligation of the RNA within the cell to create the mature mRNA that can be translated into a protein.

**[0063]** "Initiation codon" and "termination codon" refer to a unit of three adjacent nucleotides in a coding sequence that specifies initiation and chain termination respectively, of protein synthesis (mRNA translation). "Open reading frame" refers to the coding sequence uninterrupted by introns between initiation and termination codons that encodes an amino acid sequence.

**[0064]** "RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into protein by the cell. "cDNA" refers to a double-stranded DNA that is complementary to and derived from mRNA. "Sense" RNA refers to RNA transcript that includes the mRNA. "Antisense RNA" refers to a RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target gene by interfering with the processing, transport and/or translation of its primary transcript or mRNA. The complementarity of an antisense RNA may be with any part of the specific gene transcript, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. In addition, as used herein, antisense RNA may contain regions of ribozyme sequences that increase the efficacy of antisense RNA to block gene expression. "Ribozyme" refers to a catalytic RNA and includes sequence-specific endoribonucleases.

**[0065]** As used herein, "suitable regulatory sequences" refer to nucleotide sequences in native or chimeric genes that are located upstream (5'), within, and/or downstream (3') to the nucleic acid fragments of the invention, which control the expression of the nucleic acid fragments of the invention. The term "expression", as used herein, refers to the transcription and stable accumulation of the sense (mRNA) or the antisense RNA derived from the nucleic acid fragment(s) of the invention that, in conjunction with the protein apparatus of the cell, results in altered levels of the fatty acid desaturase(s). Expression or overexpression of the gene involves transcription of the gene and translation of the mRNA into precursor or mature fatty acid desaturase proteins. "Antisense inhibition" refers to the production of antisense RNA transcripts capable of preventing the expression of the target protein. "Overexpression" refers to the production of a gene product in transgenic organisms that exceeds levels of production in normal or non-transformed organisms. "Cosuppression" refers to the expression of a foreign gene which has substantial homology to an endogenous gene resulting in the suppression of expression of both the foreign and the endogenous gene. "Altered levels" refers to the production of gene product(s) in transgenic organisms in amounts or proportions that differ from that of normal or non-transformed organisms.

**[0066]** "Promoter" refers to a DNA sequence in a gene, usually upstream (5') to its coding sequence, which controls the expression of the coding sequence by providing the recognition for RNA polymerase and other factors required for proper transcription. In artificial DNA constructs promoters can also be used to transcribe antisense RNA. Promoters may also contain DNA sequences that are involved in the binding of protein factors which control the effectiveness of transcription initiation in response to physiological or developmental conditions. It may also contain enhancer elements. An "enhancer" is a DNA sequence which can stimulate promoter activity. It may be an innate element of the promoter or a heterologous element inserted to enhance the level and/or tissue-specificity of a promoter. "Constitutive promoters" refers to those that direct gene expression in all tissues and at all times. "Tissue-specific" or "development-specific" promoters as referred to herein are those that direct gene expression almost exclusively in specific tissues, such as leaves or seeds, or at specific development stages in a tissue, such as in early or late embryogenesis, respectively.

**[0067]** The "3' non-coding sequences" refers to the DNA sequence portion of a gene that contains a polyadenylation signal and any other regulatory signal capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor.

**[0068]** The term "Transit Peptide" refers to the N-terminal extension of a protein that serves as a signal for uptake and transport of that protein into an organelle such as a plastid or mitochondrion.

**[0069]** "Transformation" herein refers to the transfer of a foreign gene into the genome of a host organism and its genetically stable inheritance. "Restriction fragment length polymorphism" refers to different sized restriction fragment lengths due to altered nucleotide sequences in or around variant forms of genes. "Fertile" refers to plants that are able to propagate sexually.

**[0070]** "Oil-producing species" herein refers to plant species which produce and store triacylglycerol in specific organs, primarily in seeds. Such species include soybean (Glycine max), rapeseed and canola (including Brassica napus,

B. campestris), sunflower (Helianthus annus), cotton (Gossypium hirsutum), corn (Zea mays), cocoa (Theobroma cacao), safflower (Carthams tinctorius), oil palm (Elaeis guineensis), coconut palm (Cocos nucifera), flax (Linum usitatissimum), castor (Ricinus communis) and peanut (Arachis hypogaea). The group also includes non-agronomic species which are useful in developing appropriate expression vectors such as tobacco, rapid cycling Brassica species, and Arabidopsis thaliana, and wild species which may be a source of unique fatty acids. "Sequence-dependent protocols" refer to techniques that rely on a nucleotide sequence for their utility. Examples of sequence-dependent protocols include, but are not limited to, the methods of nucleic acid and oligomer hybridization and methods of DNA and RNA amplification such as are exemplified in various uses of the polymerase chain reaction. "PCR product" refers to the DNA product obtained through polymerase chain reaction.

**[0071]** Various solutions used in the experimental manipulations are referred to by their common names such as "SSC", "SSPE", "Denhardt's solution", etc. The composition of these solutions may be found by reference to Appendix B of Sambrook, et al. (Molecular Cloning, A Laboratory Manual, 2nd ed. (1989), Cold Spring Harbor Laboratory Press).

T-DNA Mutagenesis and Identification of an Arabidopsis Mutant Defective in Delta-15 Desaturation

**[0072]** In T-DNA mutagenesis (Feldmann, et al., Science (1989) 243:1351-1354), the integration of T-DNA in the genome can interrupt normal expression of the gene at or near the site of the integration. If the resultant mutant phenotype can be detected and shown genetically to be tightly linked to the T-DNA insertion, then the "tagged" locus and its wild type counterpart can be readily isolated by molecular cloning by one skilled in the art.

**[0073]** Arabidopsis thaliana seeds were transformed by Agrobacterium tumefaciens C58C1rif strain harboring the avirulent Ti-plasmid pGV3850::pAK1003 that has the T-DNA region between the left and right T-DNA borders replaced by the origin of replication region and ampicillin resistance gene of plasmid pBR322, a bacterial kanamycin resistance gene, and a plant kanamycin resistance gene (Feldmann, et al., Mol. Gen. Genetics (1987) 208:1-9). Plants from the treated seeds were self-fertilized and the resultant progeny seeds, germinated in the presence of kanamycin, were self-fertilized to give rise to a population, designated T3, that was segregating for T-DNA insertions. T3 seeds from approximately 6000 T2 plants were analyzed for fatty acid composition. One line, designated 3707, showed a reduced level of linolenic acid (18:3). One more round of self-fertilization of mutant line 3707 produced T4 progeny seeds. The ratio of 18:2/18:3 in seeds of the homogyzous mutant in T4 population was ca. 14; this ratio is ca 1.8 and ca. 23, respectively, in wild-type Arabidopsis and Arabidopsis fad 3 mutant [Lemieux et al. (1990) Theor. App. Gen. 80: 234-240 ] obtained via chemical mutagenesis. These seeds were planted and 263 individual plants were analyzed for the presence of nopaline in leaf extracts. T5 seeds from these plants were further analyzed for fatty acid composition and the ability to germinate in the presence of kanamycin. The mutant fatty acid phenotype was found to segregate in a 1:2:1 ratio, as was germinability on kanamycin. Nopaline was found in all plants with an altered fatty acid phenotype, but not in wild type segregants. These results provided evidence that the locus controlling delta-15 desaturation was interrupted by T-DNA in mutant line 3707.

Isolation of Arabidopsis Genomic DNA Containing the Gene Controlling Delta-15 Desaturation

**[0074]** In order to isolate the gene controlling delta-15 desaturation from wild-type Arabidopsis, a T-DNA-plant DNA "junction" fragment containing a T-DNA border integrated into the host plant DNA was isolated from Arabidopsis mutant 3707. For this, genomic DNA from the mutant plant was isolated and completely digested by either Bam HI or Sal I restriction enzymes. In each case, one of the resultant fragments was expected to contain the origin of replication and ampicillin-resistance gene of pBR322 as well as the left T-DNA-plant DNA junction fragment. Such fragments were rescued as plasmids by ligating the digested genomic DNA fragments at a dilute concentration to facilitate self-ligation and then using the ligated fragments to transform E. coli cells. Ampicillin-resistant E. coli transformants were isolated and screened by colony hybridization to fragments containing either the left or the right T-DNA border. Of the 192 colonies obtained from the plasmid rescue of Sal I digested genomic DNA, 31 hybridized with the left T-DNA border fragment, 4 hybridized to the right T-DNA border fragment, and none hybridized to both. Of the 85 colonies obtained from the plasmid rescue of Bam HI digested genomic DNA, 63 hybridized to the left border and none to the right border. Restriction analysis of seven rescued plasmids that were obtained from the Bam HI digestion and that hybridized to the left T-DNA border showed that they were indistinguishable and contained 1.4 kb of putative, flanking plant DNA. Restriction analysis of another rescued plasmid, pS1, that was obtained from the Sal I digestion and hybridized only to the left T-DNA border, showed that it contained 2.9 kb of putative, flanking plant DNA. This flanking DNA had a Bam HI site and a Hind III site 1.4 kb and 2.2 kb, respectively, away from the left T-DNA border, suggesting that the 1.4 kb putative plant DNA in Bam HI rescued plasmids was contained within the 2.9 kb putative plant DNA in the Sal I rescued plasmids. Southern blot analysis of wild type and mutant 3707 Arabidopsis genomic DNA using the radiolabeled 1.4 kb DNA fragment as the hybridization probe confirmed that this fragment contained plant DNA and that the T-DNA integration site was in a 2.8 kb Bam HI, a 5.2 kb Hind III, a 3.5 kb Sal I, a 5.5 kb Eco RI, and an approximately 9 kb

Cla I fragment of wild type Arabidopsis DNA. Nucleotide sequencing of plasmid pS1 with a primer made to a left T-DNA border sequence revealed that pS1 was colinear with the sequence of the left T-DNA border (Yadav et al., Proc. Natl. Acad. Sci. USA (1982) 79:6322-6326) up to nucleotide position 65, which is in the T-DNA border repeats. Approximately 800 bp of additional sequence in pS1 beyond the T-DNA-plant DNA junction, that is, in the plant DNA adjoining the left T-DNA border, showed no significant homology to the T-DNA of pGV3850::pAK1003 and no significant open reading frame.

**[0075]** The nucleic acid fragment from wild-type Arabidopsis corresponding to the plant DNA flanking T-DNA in the line 3707 was isolated by screening a lambda phage Arabidopsis thaliana genomic library with the 1.4 kb plant DNA isolated from the rescued plasmids as a hybridization probe. Seven positively-hybridizing genomic clones were isolated that fell in one of five classes based on partial restriction mapping. While their average insert size was approximately 15 kb, taken together they spanned a total of approximately 40 kb of genomic DNA. A combination of restriction and Southern analyses revealed that the five clones overlapped the site of integration of the left border of the T-DNA and that there was no detectable rearrangement of plant DNA in the rescued plasmids as compared to that in the wild type genomic plant DNA. One of these lambda phage clones, designated 1111, was representative of the recovered clones and contained an approximately 20 kb genomic DNA insert which was more or less symmetrically arranged around the site of insertion of the left border of the T-DNA. This clone was deposited on November 27, 1991 with the American Type Culture Collection of Rockville, Maryland under the provisions of the Budapest Treaty and bears accession number ATCC 75167.

Isolation of Arabidopsis Delta-15 Desaturase cDNA

**[0076]** A 5.2 kb Hind III fragment containing wild-type genomic DNA, which hybridized to the 1.4 kb flanking plant DNA recovered from line 3707 and which was interrupted near its middle by the T-DNA insertion in line 3707, was isolated from lambda phage clone 41A1 and cloned into the Hind III site of the pBluescript SK vector (Stratagene) by standard cloning procedures described in Sambrook et al., Molecular Cloning, A Laboratory Manual, 2nd ed. (1989), Cold Spring Harbor Laboratory Press). The resultant plasmid was designated pF1. The isolated 5.2 kb Hind III fragment was also used as a radiolabeled hybridization probe to screen a cDNA library made to poly $A^+$ mRNA from 3-day-old etiolated Arabidopsis thaliana (ecotype Columbia) seedling hypocotyls in a lambda ZAP II vector (Stratagene). Of the several positively-hybridizing plaques, four strongly-hybridizing ones were subjected to plaque purification. Sequences of the pBluescript (Stratagene) vector, including the cDNA inserts, from each of the purified phage stocks were excised in the presence of a helper phage. The resultant phagemids were used to infect E. coli cells which yielded double-stranded plasmids, pCF1, pCF2, pCF3, and pCF4. All four were shown to contain at least one approximately 1.3 to 1.4 kb Not I insert fragment (Not I/Eco RI adaptors were used in the preparation of the cDNA library) which hybridized to the same region of wild-type plant genomic DNA present in the isolated phage clones. This region, which was near the site of integration of the left T-DNA border in line 3707, was on the side of the T-DNA insertion opposite to that of the plant DNA flanking the left T-DNA border isolated previously via plasmid rescue. Partial sequence determination of the different cDNAs revealed common identity. Since multiple versions of only one type of cDNA were obtained from a cDNA library made from etiolated tissue which is expected to express delta-15 desaturation, and since these cDNAs hybridized to the genomic DNA that corresponds to the site of T-DNA integration in line 3707 which had a high linoleic acid/low linolenic acid phenotype, Applicants were lead to conclude that the T-DNA in line 3707 interrupted the normal expression of the gene encoding delta-15 desaturase. The complete nucleotide sequence of one cDNA, designated pCF3, was determined and is shown as SEQ ID NO:1. It reveals an open reading frame that encodes a 386 amino acid polypeptide. One of the sequencing primers made to the pCF3 insert was also used to obtain 255 bp of sequence from pF1 that is shown as SEQ ID NO:3. Nucleotides 68 to 255 of the genomic DNA in pF1 (SEQ ID NO:3) are identical to nucleotides 1 to 188 of the cDNA (SEQ ID NO:1), which shows that they are colinear and that the cDNA is encoded for by the gene in the isolated genomic DNA. Nucleotides 113 to 115 in SEQ ID NO:3 are the initiation codon of the largest open reading frame corresponding to nucleotides 46-48 in SEQ ID NO:1. This is evident from the presence of in-frame termination codons at nucleotides 47 to 49 and nucleotides 56 to 58 and the absence of observable intron splice junctions in SEQ ID NO:3. The identification of the 386 amino acid polypeptide as a desaturase was confirmed by comparing its amino acid sequence with all the protein sequences found in Release 19.0 of the SWISSPROTEIN database using the FASTA algorithm of Pearson and Lipman (Proc. Natl. Acad. Sci. USA (1988) 85:2444-2448) and the BLAST program (Altschul et al., J. Mol. Biol. (1990) 215:403-410). The most homologous protein found in both searches was the desA fatty acid desaturase from the cyanobacterium Synechocystis PCC6803 (Wada, et al., Nature (1990) 347:200-203; Genbank ID:CSDESA; GenBank Accession No:X53508). The 386 amino acid peptide in SEQ ID NO:1 was also compared to the 351 amino acid sequence of desA by the method of Needleman et al. (J. Mol. Biol. (1970) 48:443-453). Over their entire length, these proteins were 26% identical, the comparison imposing four major gaps in the desA protein sequence. While this overall homology is poor, homology in shorter stretches was better. For instance, in a stretch of 78 amino acids the Arabidopsis delta-15 desaturase (amino acids 78 to 155 in SEQ ID NO:1)

and the desA protein (amino acids 67 to 144) showed 40% identity and 66% similarity. Homology in yet shorter stretches was even greater as shown in Table 2.

TABLE 2

| Peptide Length | AA positions in SEO ID NO:1 | AA positions in desA | Percent Identity |
|---|---|---|---|
| 12 | 97-108 | 86-97 | 83 |
| 7 | 115-121 | 104-110 | 71 |
| 9 | 133-141 | 22-130 | 56 |
| 11 | 299-309 | 282-292 | 64 |

**[0077]** These high percent identities in short stretches of amino acids between the cyanobacterial desaturase polypeptide and SEQ ID NO:2 suggests significant relatedness between the two.

**[0078]** To analyse the developmental expression of the gene encoding mRNA coresponding to SEQ ID NO:1, the cDNA insert in plasmid pCF3 was used as a radiolabeled hybridization probe on mRNA samples from leaf, root, germinating seedling, and developing siliques from both wild type amd mutant 3707 Arabidopsis plants, essentially as described in Maniatis et al., Molecular Cloning, A Laboratory Manual (1982) Cold Spring Harbor Laboratory Press. The results indicated that while the mRNA corresponding to SEQ ID NO:1 is detected in all tissues from the mutant plant, its levels are lower than in wild-type tissues. This is consistent with the observation that the fatty acid mutation in line 3707 is leaky relative to the known Arabidopsis fad 3 mutant obtained via chemical mutagenesis. These results confirmed that the T-DNA in line 3707 had interrupted the normal expression of a fatty acid desaturase gene. Based on the fatty acid phenotype of homozygous mutant line 3707, Applicants concluded that the cDNA insert in pCF3 encoded the delta-15 desaturase. Further, Applicants concluded that it was the microsomal delta-15 desaturase, and not the chloroplastic delta-15 desaturase, since: a) the mutant phenotype was expressed strongly in the seed but expressed poorly, if at all, in the leaf of line 3707, and b) the delta-15 desaturase polypeptide, by comparison to the desA polypeptide, did not have an N-terminal extension of a transit peptide expected for a nuclear-encoded chloroplast desaturase.

**[0079]** The identity of SEQ ID NO:2 as the Arabidopsis microsomal delta-15 desaturase was confirmed by its biological overexpression in plant tissues. For this, the 1.4 kB Not I fragment of plasmid pCF3 containing the delta-15 desaturase cDNA was placed in the sense orientation behind either the CaMV 35S promotor, to provide constituitive expression, or behind the promotor for the gene encoding soybean a' subunit of the β-conglycinin (7S) seed storage protein, to provide embryo-specific expression. The chimeric genes 35S promoter/sense SEQ ID NO:1/3' nopaline synthase and β-conglycinin/sense SEQ ID NO:1/3' phaseolin were then transformed into plant cells by Agrobacterium tumefaciens's binary Ti plasmid vector system [Hoekema et al. (1983) Nature 303:179-180; Bevan (1984) Nucl. Acids Res. 12:8711-8720].

**[0080]** To confirm the identity of SEQ ID NO:1 and to test the biological effect of its overexpression in a heterologous plant species, the chimeric genes 35S promoter/sense SEQ ID NO:1/3' nopaline synthase was transformed into a binary vector, which was then transferred into Agrobacterium tumefaciens strain R1000, carrying the Ri plasmid pRiA4b from Agrobacterium rhizogenes [Moore et al. (1979) Plasmid 2:617-626]. Carrot (Daucus carota L.) cells were transformed by co-cultivation of carrot root disks with strain R1000 carrying the chimeric gene by the method of Petit et al. (1986) [Mol. Gen. Genet. 202:388-393). Fatty acid analyses of transgenic carrot "hairy" roots show that overexpression of Arabidopsis microsomal delta-15 desaturase can result in over 10-fold increase in 18:3 at the expense of 18:2.

**[0081]** To complement the delta-15 desaturation mutation in the T-DNA mutant line 3707 and to test the biological effect of overexpression of SEQ ID NO:1 in seed, the embryo-specific promoter/SEQ ID NO:1/3' phaseolin chimeric gene was transformed into a binary vector, which was then transformed into the avirulent Agrobacterium strain LBA4404/pAL4404 [Hoekema et al. (1983) Nature 303:179-180]. Roots of line 3707 were transformed by the engineered Agrobacterium, transformed plants were selected and grown to give rise to seeds. Fatty acid analysis of the seeds from two plants showed that the one out of six seeds in each plant showed the mutant fatty acid phenotype, while the remaining seeds show more than 10-fold increase in 18:3 to ca. 55%. While the sample size is small, this segregation suggests Mendelian inheritance of the fatty acid phenotype. While most of the increase occurs at the expense of 18:2, some of it also occurs at the expense of 18:1. Thus, overexpression of this gene in oils crops, especially canola, which is a close relative of Arabidopsis, is also expected to result in the high levels of 18:3 that are found in specialty oil of linseed.

**[0082]** Comparisons of the sequence of the 386 amino acid polypeptide by the method of Needleman et al. (J. Mol. Biol. (1970) 48:443-453) with those for the microsomal stearoyl-CoA (delta-9) desaturases from rat, mouse and yeast revealed 21%, 19%, and 17% identities, respectively. While the membrane-associated Arabidopsis delta-15 desaturase protein showed significant but limited homology to the desA protein, it showed no significant homology to the soluble

stearoyl-ACP (delta-9) desaturases from higher plants, including one from Arabidopsis.

**[0083]** Comparison of partial nucleotide sequences of plasmids pF1 and pS1 showed that the left T-DNA border: plant DNA junction is ca. 700 bp from the initiaton codon in SEQ ID NO:1. To determine the position of the other T-DNA: plant DNA junction with respect to the pF1 sequence, the T-DNA:plant DNA junction fragment was isolated. Genomic DNA from mutant line 3707, isolated as described previously, was partially digested by restriction enzyme Mbo I to give an average fragment size of ca. 15 kB. The fragment ends were partially-filled with dGTP and gATP by Klenow and cloned into Xho I half-sites of LambdaGEM®-11 (Promega Corporation) following the manufacturer's protocol. The phage library was titered and used essentially as described in Ausubel et al. [Current Protocols in Molecular Biology (1989) John Wiley & Sons]. The genomic phage library was screened with radiolabeled PCR product, ca. 0.6 kB, derived from 5' end of the gene in pF1. This product spans from 3 bp to the right of where the left-T-DNA border inserted to 15 bp to the left of nucleotide position 1 in SEQ ID NO:1. Southern blot analysis of DNA from one of the purified, positively-hybridizing phages following Eco RI restriction digestion and electrophoresis showed that a 4 kB Eco RI fragment hybridized to the 0.6 kB PCR product. The Eco RI fragment was subcloned and subject to sequence analyses. Comparison of the sequences derived from this fragment, pF1 and pS1 showed that the insertion of T-DNA resulted in a 56 bp deletion at the site of insertion and that the T-DNA interrupted the Arabidopsis gene 711 bp 5' to the initiaton codon in SEQ ID NO:1. Thus, the T-DNA inserts 5' to the open reading frame, consistent with the leaky expresssion of the gene encoding SEQ ID NO:1 and the leaky fatty acid phenotype in mutant 3707. While the left T-DNA:plant DNA junction is precise, that is without any sequence rearrangement in either the left T-DNA border or the flanking plant DNA, the other T-DNA:plant DNA junction is complex and not fully characterized.

**[0084]** Plasmid pCF3 was deposited on December 3, 1991 with the American Type Culture Collection of Rockville, Maryland under the provisions of the Budapest Treaty and bears accession number ATCC 68875.

Using Arabidopsis Delta-15 Desaturase cDNA as a Hybridization Probe to Isolate cDNAs Encoding Related Desaturases from Arabidopsis

**[0085]** The 1.4 kb Not I insert fragment isolated from plasmid pCF3 was purified, radiolabeled, and used to screen approximately 80,000 clones from the cDNA library made to poly $A^+$ mRNA from 3-day-old etiolated Arabidopsis thaliana as described above, except that lower stringency hybridizations (1 M NaCl, 50 mM Tris-HCl, pH 7.5, 1% SDS, 5% dextran sulfate, 0.1 mg/mL denatured salmon sperm DNA and 50°C) and washes (sequentially with 2X SSPE, 0.1% SDS at room temperature for 5 min and then again with fresh solution for 10 min, and finally with 0.5X SSPE, 0.1% SDS at 50°C for 5 min.) were used. Approximately 17 strongly-hybridizing and 17 weakly-hybridizing plaques were identified in the primary screen. Four of the weakly-hybridizing plaques were picked and subjected to one or two further rounds of screening with the radiolabeled probe as above until they were pure. To ensure that these were not delta-15 desaturase clones, they were further analyzed to determine whether they hybridized to an 18 bp oligomer specific to the 3' non-coding region of delta-15 desaturase cDNA (pCF3). After autoradiography of the filters, one of the clones was found not to hybridize to this probe. This clone was picked, and a plasmid clone containing the cDNA insert was obtained as described above. Restriction analysis of this plasmid, designated pCM2, showed that it had an approximately 1.3 kb cDNA insert which lacked a 0.7 kb Nco I - Bgl II fragment characteristic of the Arabidopsis delta-15 desaturase cDNA of pCF3. (This fragment corresponds to the DNA located between the Nco I site at nucleotides 474 to 479 and the Bgl II site at nucleotides 1164 to 1169 in SEQ ID NO:1). Partial nucleotide sequences of single strands from the 5' region and 3' region of pCM2 revealed that the cDNA insert was incomplete and that it encoded a polypeptide that is similar to, but distinct from, that encoded by the cDNA in pCF3. In order to isolate a full-length version of the cDNA in plasmid pCM2, the 1.3 kB Not I fragment from plasmid pCM2 containing the cDNA insert was isolated and used as a radiolabeled hybridization probe to rescreen the same Arabidopsis cDNA library as above. Three strongly hybridizing plaques were purified and the plasmids excised as described previously. The three resultant plasmids were digested by Not I restriction enzyme and shown to contain cDNA inserts ranging in size between 1 kB and 1.5 kB. Complete nucleotide sequence determination of the cDNA insert in one of these plasmids, designated pACF2-2, is shown in SEQ ID NO:4. SEQ ID NO:4 shows the 5' to 3' nucleotide sequence of base pairs of the Arabidopsis thaliana cDNA which encodes a fatty acid desaturase. Nucleotides 10-12 and nucleotides 1358 to 1350 are, respectively, the putative initiation codon and the termination codon of the open reading frame (nucleotides 10 to 1350). The open reading frame was confirmed by comparison of its deduced amino acid sequences with that of the related delta-15 fatty acid desaturase from soybean in this application. Nucleotides 1 to 9 and 1351 to 1525 are, respectively, the 5' and 3' untranslated nucleotides. The 446 amino acid protein sequence in SEQ ID NO:5 is that deduced from the open reading frame in SEQ ID NO:4 and has an estimated molecular weight of 51 kD. Alignment of SEQ ID NOS:2 and 5 shows an overall homology of approximately 80% and that the former has an approximately 55 amino acid long N-terminal extension, which is deduced to be a transit peptide found in nuclear-encoded plastid proteins.

**[0086]** To analyse the developmental expression of the gene corresponding to SEQ ID NO:4, this sequence was used as a radiolabeled hybridization probe on mRNA samples from leaf, root, germinating seedling, and developing

siliques from both wild type and mutant line 3707 Arabidopsis plants, essentially as described in Maniatis et al. [Molecular Cloning, A Laboratory Manual (1982) Cold Spring Harbor Laboratory Press]. The results indicated that, in contrast to the constitutive expression of the gene encoding SEQ ID NO:1, the mRNA corresponding to SEQ ID NO:4 is abundant in green tissues, rare in roots and leaves, and is about threefold more abundant in leaf than that of SEQ ID NO:1. The cDNA in plasmid pCM2 was also shown to hybridize polymorphically to genomic DNA from Arabidopsis thaliana (ecotype Wassileskija and marker line W100 ecotype Landesberg background) digested with Eco RI. It was used as a RFLP marker to map the genetic locus for the gene encoding this fatty acid desaturase in Arabidopsis. A single genetic locus was positioned corresponding to this desaturase cDNA. Its location was thus determined to be on chromosome 3 between the lambda AT228 and cosmid c3838 RFLP markers, "north" of the glabrous locus (Chang et al., Proc. Natl. Acad. Sci. USA (1988) 85:6856-6860; Nam et al., Plant Cell (1989) 1:699-705). This approximates the region to which Arabidopsis fatty acid desaturase fad 2, fad D, and fad B mutations map [Somerville et al., (1992) in press]. Unsuccessful efforts to clone the microsomal delta-12 fatty acid desaturase using cDNA inserts of pCF3 and pACF2-2 alongwith the above data led Applicants to conclude that the cDNA in pACF2-2 encodes a plastid delta-15 fatty acid desaturase that corresponds to the fad D locus. This conclusion will be confirmed by biological expression of the cDNA in pACF2-2.

**[0087]** Plasmid pCM2 was deposited on November 27, 1991 with the American Type Culture Collection of Rockville, Maryland under the provisions of the Budapest Treaty and bears accession number ATCC 68852.

**[0088]** The 1.4 kb, 1.3 kB, and 1.5 kB Not I cDNA insert fragments isolated from plasmids pCF3, pCM2 and pACF2-2 were purified, radiolabeled, and used several times to screen at low stringency as described above two different cDNA libraries: one was made to poly $A^+$ mRNA from 3-day-old etiolated Arabidopsis thaliana ("etiolated" library) as described above and one made to poly$A^+$ mRNA from the above-ground parts of Arabidopsis thaliana plants, which varied in size from those that had just opened their primary leaves to plants which had bolted and were flowering [Elledge et al. (1991) Proc. Natl. Acad Sci. USA 88:1731-1735]. The cDNA inserts in the library were made into an Xho I site flanked by Eco RI sites in lambda Yes vector [Elledge et al. (1991) Proc. Natl. Acad Sci. USA 88:1731-1735] ("leaf" library). Several plaques from both libraries that hybridized weakly and in duplicate lifts to both SEQ ID NOS:1 and 4 were subjected to plaque purification. Phagemids were excised from the pure phages from "etiolated" library as described above. Plasmids were excised from the purified phages of the "leaf" library by site-specific recombination using the cre-lox recombination system in E. coli strain BNN132 [Elledge et al. (1991) Proc. Natl. Acad Sci. USA 88:1731-1735]. In all cases, nucleotide sequencing of the cloned DNA revealed clones either identical to SEQ ID NOS:1 or 4 or unrecognizable sequences.

**[0089]** In another set of experiments ca. 400,000 phages in the "leaf" library was screened with SEQ ID NOS:1 and 4 at low stringency (26 C, 1 M Na+, 50% formamide) and high stringency (42 C, 1 M Na+, 50% formamide). Of the several positive signals on the primary plaque lifts, 11 showed high stringency hybridization to SEQ ID NO:1, 35 showed high stringency hybridization to SEQ ID NO:4, and 39 hybridized to both at low stringency only. Twenty seven plaques of the low stringency signals came through a secondary low-stringency screen, 17 of which were used to make DNA from excised plasmids. Of the 7 plasmid DNA were sequenced, 8 were unrecognizable sequences, 5 were identical to SEQ ID NO:1, 2 were identical to SEQ ID NO:2, and 2 were identical to one another and related but distinct to SEQ ID NOS:1 and 4. The novel desaturase sequence, designated pFad-x2, was also isolated from the "leaf" library independently by using as a hybridization probe a 0.6 kB PCR product derived by polymerase chain reaction on poly A+ RNA made from both canola seed as well as Arabidopsis leaves, as described elsewhere in this application, using degenerate oligomers made to conserved sequences between plant delta-15 desaturases and the cyanobacterial des A desaturase. The PCR-derived plasmid, designated pYacp7, was sequenced partially from both ends. Comparison of the sequences of pFad-x2 and pYacp7 revealed that the two independently cloned cDNAs contained an identical sequence that was related to the other delta-15 desaturases and that both were incomplete cDNAs. A partial composite sequence derived from both plasmids, pFadx-2 and pYacp7, is shown in SEQ ID NO:16 as a 5' to 3' nucleotide sequence of 472 bp. Nucleotides 2-4 and nucleotides 468 to 470 are, respectively, the first and the last codons in the open reading frame. This open reading frame is shown in SEQ ID NO:17. Comparison of SEQ ID NO:17 to the other delta-15 desaturase polypeptides disclosed in this application by the method of Needleman et al. [J. Mol. Biol. (1970) 48:443-453)] using gap weight and gap length weight values of 3.0 and 0.1, respectively. The overall identities are between 65% and 68% between SEQ ID NO:17 and the microsomal delta-15 desaturases from Arabidopsis, canola and soybean and the overall identities are between 77% and 87% between SEQ ID NO:17 and the plastid delta-15 desaturases from Arabidopsis, canola and soybean. In addition SEQ ID NO:17 has an N-terminal peptide extension compared to the microsomal delta-15 desaturases that shows homology of the transit peptide sequence in Arabidopsis plastid delta-15 desaturase. On the basis of these comparisons it is deduced that SEQ ID NO:16 encodes a plastid delta-15 desaturase. There is genetic data in Arabidopsis suggesting the presence of two loci for plastid delta-15 desaturase. The full-length version of SEQ ID NO:16 can be readily isolated by one skilled in the art. The biological effect of introducing SEQ ID NO:16 or its full-length version into plants will be used to confirm its identity.

**[0090]** Plasmid pYacp7 was deposited on 20 November 1992 with the American Type Culture Collection of Rockville,

Maryland under the provisions of the Budapest Treaty and bears accession number ATCC 69129.

Using Arabidopsis Delta-15 Desaturase cDNAs as Hybridization Probes to Isolate Delta-15 Desaturase cDNAs from Other Plant Species

**[0091]** For the purpose of cloning the Brassica napus seed cDNAs encoding delta-15 fatty acid desaturases, the cDNA inserts from pCF3 and pCM2 were isolated by polymerase chain reaction from the respective plasmids, radiolabeled, and used as hybridization probes to screen a lambda phage cDNA library made with poly $A^+$ mRNA from developing Brassica napus seeds 20-21 days after pollination. This cDNA library was screened several times at low stringency, using the Arabidopsis cDNA probes mentioned above. One of the Brasssica napus cDNAs obtained in the initial screens was used as probe in a subsequent high stringency screen.

**[0092]** Arabidopsis pCM2 insert was radiolabeled and used as probe to screen approximately 300,000 plaques under low stringency hybridization conditions. The filter hybridizations were performed in 50 mM Tris pH 7.6, 6X SSC, 5X Denhardt's, 0.5% SDS, 100 ug denatured calf thymus DNA at 50°C overnight, and the posthybridization washes were carried out in 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45°C for 30 min, and then repeated twice with 0.2X SSC, 0.5% SDS at 50°C for 30 min. Five strongly-hybridizing phages were obtained. These were plaque purified and used to excise the phagemids as described in the manual of the pBluescriptII Phagemid Kit from Stratagene (Stratagene 1991 catalogue, item 212205). One of these, designated pBNSF3-2, contained a 1.3 kb insert. pBNSF3-f2 was sequenced completely on both strands and the nucleotide sequence is shown in SEQ ID NO:6. Plasmid pBNSF3-2 was deposited on 27 November 1991 with the American Type Culture Collection of Rockville Maryland, USA under the provisions of the Budapest Treaty and bears the accession number 68854.

**[0093]** An additional low stringency screen using pCM2 probe provided eight strongly hybridizing phages. One of these, designated pBNSFd 8, contained a 0.4kb insert. pBNSFd-8 was sequenced completely on one strand, this nucleotide sequence showed significant divergence from the sequence SEQ ID NO:6 in the homologous region, which suggested that it corresponded to a novel Brassica napus seed desaturase different from that shown in SEQ ID NO: 6. pBNSFd-8 insert was radiolabelled and used as hybridization probe in a high stringency screen of the Brassica napus seed cDNA library. The hybridization conditions were identical to those of the low stringency screen described above except for the temperature of the final two 30 min posthybridization washes in 0.2x SSC, 0.5% SDS was increased to 60°C. This screen resulted in three strongly hybridizing phages that were purified and excised. One of the excised plasmids pBNSFd-3 contained a 1.4kb insert that was sequenced completely on both strands. SEQ ID NO:8 shows the complete nucleotide sequence of pBNSFd-2.

Using Arabidopsis Delta-15 Desaturase cDNA as a Hybridization Probe to Isolate a Glycerolipid Desaturase cDNA from Soybean

**[0094]** A cDNA library was made to poly $A^+$ mRNA isolated from developing soybean seeds, and screened essentially as described above, except that filters were prehybridized in 25 mL of hybridization buffer consisting of 50mM Tris-HCl, pH 7.5, 1 M NaCl, 1% SDS, 5% dextran sulfate and 0.1 mg/mL denatured salmon sperm DNA (Sigma Chemical Co.) at 50°C for 2 h. Radiolabeled probe prepared from pCF3 as described above was added, and allowed to hybridize for 18 h at 50°C. The probes were washed twice at room temperature with 2X SSPE, 1% SDS for five min followed by washing for 5 min at 50°C in 0.2X SSPE, 1% SDS. Autoradiography of the filters indicated that there was one strongly hybridizing plaque, and approximately five weakly hybridizing plaques. The more strongly hybridizing plaque was subjected to a second round of screening as before, except that the final wash was for 5 min at 60°C in 0.2X SSPE, 1% SDS. Numerous, strongly hybridizing plaques were observed, and one, well-isolated from other phage, was picked for further analysis.

**[0095]** Sequences of the pBluescript vector from the purified phage, including the cDNA insert, were excised in the presence of a helper phage and the resultant phagemid was used to infect E. coli XL-1 Blue cells. DNA from the plasmid, designated pXF1, was made by the alkaline lysis miniprep procedure described in Sambrook et al. (Molecular Cloning, A Laboratory Manual, 2nd ed. (1989) Cold Spring Harbor Laboratory Press). The alkali-denatured double-stranded DNA from pXF1 was completely sequenced on both strands. The insert of pXF1 contained a stretch of 1783 nucleotides which contained an unknown open-reading frame and also contained a poly-A stretch of 16 nucleotides 3' to the open reading frame, from nucleotides 1767 to 1783, followed by an Eco RI restriction site. The 2184 bases that followed this Eco RI site contained a 1145 bp open reading frame which encoded a polypeptide of about 68% identity to, and colinear with, the Arabidopsis delta-15 desaturase polypeptide listed in SEQ ID No:2. The putative start methionine of the 1145 bp open-reading frame corresponded to the start methionine of the Arabidopsis microsomal delta-15 peptide and there were no amino acids corresponding to a plastid transit peptide 5' to this methionine. When the insert in pXF1 was digested with Eco RI four fragments were observed, fragments of approximately 370 bp and 1400 bp fragments, derived from the first 1783 bp of the insert in pXF1, and fragments of approximately 600 bp and

1600 bp derived from the the other 2184 nucleotides of the insert in pXF1. Only the 600 bp and 1600 bp fragments hybridized with probe derived from pCF3 on Southern blots. It was deduced that pXF1 contained two different cDNA inserts separated by an Eco RI site and the second of these inserts was a 2184 bp cDNA encoding a soybean microsomal delta-15 desaturase. The complete nucleotide sequence of the 2184 bp soybean microsomal delta-15 cDNA contained in plasmid pXF1 is listed in SEQ ID No:10. Plasmid pXF1 was deposited on December 3, 1991 with the American Type Culture Collection of Rockville, Maryland under the provisions of the Budapest Treaty and bears accession number ATCC 68874.

Using Soybean Microsomal Delta-15 Desaturase cDNA as a Hybridization Probe to Isolate cDNAS Encoding Related Desaturases from Soybean

**[0096]** A 1.0 kb fragment of DNA corresponding to part of the coding region of the soybean microsomal delta-15 desaturase cDNA contained in plasmid pXF1, was excised with the restriction enzyme Hha I and gel purified. The fragment was labeled with $^{32}P$ as described above and used to probe a soybean cDNA library as described above. Autoradiography of the filters indicated that there were eight hybridizing plaques and these were subjected to a second round of screening. Sequences of the pBluescript vector from all eight of the purified phages, including the cDNA inserts, were excised in the presence of a helper phage and the resultant phagemids were used to infect E. coli XL-1 Blue cells. DNA from the plasmids was made by the alkaline lysis miniprep procedure described in Sambrook et al. (Molecular Cloning, A Laboratory Manual, 2nd ed. (1989) Cold Spring Harbor Laboratory Press). Restriction analysis showed they contained inserts ranging from 1.0 kb to 3.0 kb in size. One of these inserts, designated pSFD-118bwp, contained an insert of about 1700 bp. The alkali-denatured double-stranded DNA from pSFD-118bwp was completely sequenced on both strands, shown in SEQ ID NO:12. The insert of pSFD-118bwp contained a stretch of 1675 nucleotides which contained an open-reading frame encoding a polypeptide, shown in SEQ ID NO:13, of about 80% identity with, and colinear with, the Arabidopsis plastid delta-15 desaturase polypeptide listed in SEQ ID No:5. The open-reading frame also encoded amino acids corresponding to a plastid transit peptide at the 5' end of the open-reading frame. The transit peptide was colinear with, and shared some homology to, the transit peptide described for the Arabidopsis plastid delta-15 glycerolipid desaturase. The complete nucleotide sequence of the 1675 bp soybean plastid delta-15 glycerolipid desaturase cDNA is listed in SEQ ID No:12.

**[0097]** Comparison of the different delta-15 desaturase sequences disclosed in the application by the method of Needleman et al. (J. Mol. Biol. (1970) 48:443-453) using gap weight and gap length weight values of 3.0 and 0.1, respectively, reveals the relatedness between them as shown in Table 3.

TABLE 3

| Percent Identities Between Different Delta-15 Fatty Acid Desaturases at the Amino Acid Level | | | | | |
|---|---|---|---|---|---|
| | aD | c3 | cD | s3 | sD |
| a3 | 66 | 93 | 66 | 68 | 67 |
| aD | - | 67 | 90 | 67 | 69 |
| c3 | - | - | 68 | 68 | 68 |
| cD | - | - | - | 68 | 74 |

**[0098]** a3, ad, c3, cD, s3 and sD refer, respectively, to SEQ ID NO:2 (Arabidopsis microsomal delta-15 desaturase), SEQ ID NO:5 (Arabidopsis plastid delta-15 desaturase), SEQ ID NO:7 (canola microsomal delta-15 desaturase), SEQ ID NO:9 (canola plastid delta-15 desaturase), SEQ ID NO:11 (soybean microsomal delta-15 desaturase), and SEQ ID NO:13 (soybean plastid delta-15 desaturase). Based on these comparisons, the delta-15 desaturases, of both microsomal and plastid types, have overall identities of 65% or more at the amino acid levels, even when from different plant species.

Isolation of Nucleotide Sequences Encoding Homologous and Heterologous Glycerolipid Desaturases

**[0099]** Fragments of the instant invention may be used to isolate cDNAs and genes of homologous and heterologous glycerolipid desaturases from the same species as the fragment of the invention or from different species. Isolation of homologous genes using sequence-dependent protocols is well-known in the art. Southern blot analysis revealed that the Arabidopsis microsomal delta-15 desaturase cDNA (SEQ ID NO:1) hybridized to genomic DNA fragments of corn and soybean. In addition, Applicants have demonstrated that it can be used to isolate cDNAs encoding seed microsomal delta-15 desaturases from Brassica napus (SEQ ID NO:6) and soybean (SEQ ID NO:10). Thus, one can isolate cDNAs and genes for homologous glycerolipid desaturases from the same or different higher plant species, especially from

the oil-producing species.

**[0100]** More importantly, one can use the fragments of the invention to isolate cDNAs and genes for heterologous glycerolipid desaturases, including those found in plastids. Thus, Arabidopsis microsomal delta-15 desaturase cDNA (SEQ ID NO:1) was successfully used as a hybridization probe to isolate cDNAs encoding the related plastid delta-15 desaturases from Arabidopsis (SEQ ID NO:4) and Brassica napus (SEQ ID NO: 8), and the soybean microsomal delta-15 soybean (SEQ ID NO:10) was successfully used to isolate soybean cDNA encoding plastid delta-15 desaturase (SEQ ID NO:12).

**[0101]** In a particular embodiment of the present invention, regions of the nucleic acid fragments of the invention that are conserved between different desaturases may be used by one skilled in the art to design a mixture of degenerate oligomers for use in sequence-dependent protocols aimed at isolating nucleic acid fragments encoding other homologous or heterologous glycerolipid desaturase cDNA's or genes. For example, by comparing all desaturase polypeptides one can identify stretches of amino acids that are conserved between them, and then use the conserved amino acid sequence to design oligomers, both short degenerate or long ones, or "guessmers" as known by one skilled in the art (see Sambrook et al., (Molecular Cloning, A Laboratory Manual, 2nd ed. (1989), Cold Spring Harbor Laboratory Press). Such oligomers and "quessmers" may be used as hybridization probes as known to one skilled in the art.

**[0102]** For example, comparison of cyanobacterial desA and plant delta-15 desaturases revealed a particularly well conserved stretch of amino acids (amino acids 97-108 in SEQ ID NO:1). SEQ ID NOS:20 and 21 represent two sets of 36-mers each 16-fold degenerate made to this region. End-labeled oligomers represented in SEQ ID NOS:20 and 21 were mixed and used as hybridzation probes to screen Arabidopsis cDNA libraries. Most of the positively-hybridizing plaques also hybridized to cDNAs encoding Arabidopsis microsomal and plastid delta-15 desaturases (SEQ ID NOS: 1 and 4). However, the use of SEQ ID NOS:20 and 21 did not give consistent and reproducible results. A 135 base-long oligomer (SEQ ID NO:32) was also made as an antisense strand to a longer stretch of the same conserved region, amino acids 97 to 141 in SEQ ID NO:1 (FVLGHDCGHGSFSDIPLLNSVVGHILHSFILVPYHGWRISHRTHH). At positions of ambiguity, the design used either deoxyinosines or most frequently used codons based on the codon usage in Arabidopsis genes. When used as a hybridization probe, the 135-mer hybridized to all plaques that also hybridized to cDNAs encoding Arabidopsis microsomal and plastid delta-15 desaturases (SEQ ID NOS:1 and 4). In addition, it also hybridized to plaques that did not hybridize to SEQ ID NOS:1 and 4). The latter were purified and excised as described previously. Nucleotide sequencing of the cDNA inserts in the resultant plasmids revealed DNA sequences that did not show any relatedness to any desaturase.

**[0103]** For another example, in the polymerase chain reaction (Innis, et al., Eds, (1990) PCR Protocols: A Guide to Methods and Applications, Academic Press, San Diego), two short pieces of the present fragment of the invention can be used to amplify a longer glycerolipid desaturase DNA fragment from DNA or RNA. The polymerase chain reaction may also be performed on a library of cloned nucleotide sequences with one primer based on the fragment of the invention and the other on either the poly $A^+$ tail or a vector sequence. These oligomers may be unique sequences or degenerate sequences derived from the nucleic acid fragments of the invention. The longer piece of homologous glycerolipid desaturase DNA generated by this method could then be used as a probe for isolating related glycerolipid desaturase genes or cDNAs from Arabidopsis or other species. The design of oligomers, including long oligomers using deoxyinosine, and "guessmers" for hybridization or for the polymerase chain reaction are known to one skilled in the art and discussed in Sambrook et al., (Molecular Cloning, A Laboratory Manual, 2nd ed. (1989), Cold Spring Harbor Laboratory Press). Stretches of conserved amino acids between delta-15 desaturase and other desaturases, especially desA, allow for the design of such oligomers. For example, conserved stretches of amino acids between desA and delta-15 desaturase, discussed above, are useful in designing long oligomers for hybridization as well as shorter ones for use as primers in the polymerase chain reaction. In this regard, the conserved amino acid stretch of amino acids 97 to 108 of SEQ ID NO:2 is particularly useful. Other conserved regions in SEQ ID NO:2 useful for this purpose are amino acids 299 to 309, amino acids 115 to 121, and amino acids 133 to 141. Amino acid stretch 133 to 141 in SEQ ID NO:2 shows especially good homology to several desaturases. For example, in this stretch, amino acids 133, 137, 138, 140 and 141 are conserved in plant delta-15 desaturases, cyanobacterial desA, yeast and mammalian microsomal stearoyl-CoA desaturases. Comparison of cyanobacterial des A and plant delta-15 desaturases revealed two particularly well conserved stretch of amino acids (amino acids 97-108 and amino acids 299-311 in SEQ ID NO: 1) that can be used for PCR. The following sets of PCR primers were made to these regions:

| SEQ ID NO | Length | Fold Degeneracy | AA positions in SEQ ID NO:2 | AA Sequence |
|---|---|---|---|---|
| 20 | 36 | 16 | 97-108 (S) | FVLGHDCGHGSF |
| 21 | 36 | 16 | 97-108 (S) | FVLGHDCGHGSF |
| 28 | 36 | 16 | 97-108 (S) | FVLGHDCGHGSF |
| 29 | 36 | 16 | 97-108 (S) | FVLGHDCGHGSF |
| 22 | 18 | 72 | 100-105 (S) | GHDCGH |
| 23 | 18 | 72 | 100-105 (S) | GHDCGH |
| 24 | 18 | 72 | 299-304 (AS) | HDIGTH |
| 25 | 18 | 72 | 299-304 (AS) | HDIGTH |
| 26 | 23 | 416 | 304-309 (AS) | HVIHHL |
| 27 | 23 | 416 | 304-309 (AS) | HVIHHL |
| 30 | 38 | 64 | 299-311 (AS) | HDIGTHVIHHLFP |
| 31 | 38 | 64 | 299-311 (AS) | HDIGTHVIHHLFP |

**[0104]** In one experiment, PCRs were performed using SEQ ID NOS:22 and 23 as sense primers and either SEQ ID NOS:24 and 25 or SEQ ID NOS:26 and 27 as antisense primers on poly A+ RNA purified from both Arabidopsis leaf and canola developing seeds. All PCRs resulted in PCR products of the correct size (ca. 630 bp). The PCR products from Arabidopsis and canola were purified and used as radiolabeled hybridization probes to screen the Lambda Yes Arabidopsis cDNA library, as described above. This led to the isolation of a pure phage, which was excised to give plasmid pYacp7. The cDNA insert in pYacp7 was partially sequenced. It's sequence showed that it encoded an incomplete desaturase polypeptide that was identical to another cDNA (in plasmid pFadx-2) isolated by low-stringency hybridization as described previously. The composite sequence derived from the partial sequences from the cDNA inserts in pFadx-2 and pYacp7 is shown in SEQ ID NO:16 and the polypeptide encoded by it in SEQ ID NO:17. As discussed previously, SEQ ID NO:17 is a putative plastid delta-15 desaturase. This is further supported by Southern blot analysis using radiolabeled cDNA inserts from either pCF3, pACF2-2, or pYacp7 on Arabidopsis genomic DNA digested with one of several enzymes. It shows that the different inserts hybridize to different restriction fragments and that only the inserts from pACF2-2 and pYacp7 show some cross-hybridization.

**[0105]** In another PCR experiment, PCR was performed using ca. 80 pmoles each of SEQ ID NOS:28 and 29 as sense primers and ca. 94 pmoles each of SEQ ID NOS:30 and 31 as antisense primers on poly A+ RNA purified from Arabidopsis mutant line 3707. This was performed using GeneAmp® RNA PCR Kit (Perkin Elmer Cetus) following manufacturer's protocol and using the following program: a) 1 cycle of 2 min at 95°C, b) 35 cycles of 1 min at 95°C (denaturation), 1 min at 50°C (annealing) and 1 min at 65°C (extension), and c) 1 cycle of 7 min at 65°C. The resulting PCR product, of the correct size (ca. 630 bp), was purified, radiolabeled, and used as a hybridization probe on a Southern blot of Arabidopsis genomic DNA as described above. While it hybridized to restriction fragments that also hybridized to SEQ ID NOS:1 (Arabidopsis microsomal delta-15 desaturase), 4 (Arabidopsis plastid delta-15 desaturase), and 16 (Arabidopsis plastid delta-15 desaturase), it also hybridized to novel fragments that did not hybridze to previously cloned desaturase cDNAs. However, even after several attempts, the radiolabeled PCR product did not hybridize to any novel cDNA clone when used as a probe on different Arabidopsis cDNA libraries: in all cases it hybridzed only to plaques that also hybridized to the known desaturase cDNAs. Furthermore, the PCR product was subcloned into a plasmid vector and after screening about a 100 of these, none gave rise to a clone with a novel desaturase sequence.

**[0106]** The isolation of other glycerolipid desaturases will become easier as more examples of glycerolipid desaturases are isolated using the fragments of the invention. Knowing the conserved amino acid sequences from diverse desaturases will also allow one to identify more and better consensus sequences. Such sequences can be used to make hybridization probes or amplification primers which will further aid in the isolation of different glycerolipid desaturases, including those from non-plant sources such as fungi, algae, and even cyanobacteria, as well as other membrane-associated desaturases from other organisms.

**[0107]** The function of the diverse nucleotide fragments encoding glycerolipid desaturases that can be isolated using

the present invention can be identified by transforming plants with the isolated desaturase sequences, linked in sense or antisense orientation to suitable regulatory sequences required for plant expression, and observing the fatty acid phenotype of the resulting transgenic plants. Preferred target plants for the transformation are the same as the source of the isolated nucleotide fragments when the goal is to obtain inhibition of the corresponding endogenous gene by antisense inhibition or cosuppression. Preferred target plants for use in expression or overexpression of the isolated nucleic acid fragments are plants with known mutations in desaturation reactions, such as the Arabidopsis desaturase mutants, mutant flax deficient in delta-15 desaturation, or mutant sunflower deficient in delta-12 desaturation. Alternatively, the function of the isolated nucleic acid fragments can be determined similarly via transformation of other organisms, such as yeast or cyanobacteria, with chimeric genes containing the nucleic acid fragment and suitable regulatory sequences followed by analysis of fatty acid composition and/or enzyme activity.

Overexpression of the Glycerolipid Desaturase Enzymes in Transgenic Species

**[0108]** The nucleic acid fragment(s) of the instant invention encoding functional glycerolipid desaturase(s), with suitable regulatory sequences, can be used to overexpress the enzyme(s) in transgenic organisms. Such recombinant DNA constructs may include either the native glycerolipid desaturase gene or a chimeric glycerolipid desaturase gene isolated from the same or a different species as the host organism. For overexpression of glycerolipid desaturase(s), it is preferable that the introduced gene be from a different species to reduce the likelihood of cosuppression. For example, overexpression of delta-15 desaturase in soybean, rapeseed, or other oil-producing species to produce altered levels of polyunsaturated fatty acids may be achieved by expressing RNA from the entire cDNA found in pCF3. Similarly, the isolated nucleic acid fragments encoding glycerolipid desaturases from Arabidopsis, rapeseed, and soybean can also be used by one skilled in the art to obtain substantially homologous full-length cDNAs, if not already obtained, as well as the corresponding genes as fragments of the invention. These, in turn, may be used to overexpress the corresponding desaturases in plants. One skilled in the art can also isolate the coding sequence(s) from the fragment(s) of the invention by using and/or creating sites for restriction endonucleases, as described in Sambrook et al., (Molecular Cloning, A Laboratory Manual, 2nd ed. (1989), Cold Spring Harbor Laboratory Press). For example, the fragment in SEQ ID NO:1 in plasmid pCF3 is flanked by Not I sites and can be isolated as a Not I fragment that can be introduced in the sense orientation relative to suitable plant regulatory sequences. Alternatively, sites for Nco I (5'-CCATGG-3') or Sph I (5'-GCATGC-3') that allow precise removal of coding sequences starting with the initiating codon "ATG" may be engineered into the fragment(s) of the invention. For example, for utilizing the coding sequence of delta-15 desaturase from pCF3, an Sph I site can be engineered by substituting nucleotides at positions 44, 45, and 49 of SEQ ID NO:1 with G, C, and C, respectively.

Inhibition of Plant Target Genes by Use of Antisense RNA

**[0109]** Antisense RNA has been used to inhibit plant target genes in a tissue-specific manner (see van der Krol et al., Biotechniques (1988) 6:958-976). Antisense inhibition has been shown using the entire cDNA sequence (Sheehy et al., Proc. Natl. Acad. Sci. USA (1988) 85:8805-8809) as well as a partial cDNA sequence (Cannon et al., Plant Molec. Biol. (1990) 15:39-47). There is also evidence that the 3' non-coding sequences (Ch'ng et al., Proc. Natl. Acad. Sci. USA (1989) 86:10006-10010) and fragments of 5' coding sequence, containing as few as 41 base-pairs of a 1.87 kb cDNA (Cannon et al., Plant Molec. Biol. (1990) 15:39-47), can play important roles in antisense inhibition.

**[0110]** The use of antisense inhibition of the glycerolipid desaturases may require isolation of the transcribed sequence for one or more target glycerolipid desaturase genes that are expressed in the target tissue of the target plant. The genes that are most highly expressed are the best targets for antisense inhibition. These genes may be identified by determining their levels of transcription by techniques, such as quantitative analysis of mRNA levels or nuclear run-off transcription, known to one skilled in the art.

**[0111]** For example, antisense inhibition of delta-15 desaturase in Brassica napus resulting in altered levels of polyunsaturated fatty acids may be achieved by expressing antisense RNA from the entire or partial cDNA found in pBNSF3-2.

Inhibition of Plant Target Genes by Cosuppression

**[0112]** The phenomenon of cosuppression has also been used to inhibit plant target genes in a tissue-specific manner. Cosuppression of an endogenous gene using the entire cDNA sequence (Napoli et al., The Plant Cell (1990) 2: 279-289; van der Krol et al., The Plant Cell (1990) 2:291-299) as well as a partial cDNA sequence (730 bp of a 1770 bp cDNA) (Smith et al., Mol. Gen. Genetics (1990) 224:477-481) are known.

**[0113]** The nucleic acid fragments of the instant invention encoding glycerolipid desaturases, or parts thereof, with suitable regulatory sequences, can be used to reduce the level of glycerolipid desaturases, thereby altering fatty acid

composition, in transgenic plants which contain an endogenous gene substantially homologous to the introduced nucleic acid fragment. The experimental procedures necessary for this are similar to those described above for the overexpression of the glycerolipid desaturase nucleic acid fragments. For example, cosuppression of delta-15 desaturase in Brassica napus resulting in altered levels of polyunsaturated fatty acids may be achieved by expressing in the sense orientation the entire or partial seed delta-15 desaturase cDNA found in pBNSF3-2.

Selection of Hosts, Promoters and Enhancers

**[0114]** A preferred class of heterologous hosts for the expression of the nucleic acid fragments of the invention are eukaryotic hosts, particularly the cells of higher plants. Particularly preferred among the higher plants are the oil-producing species, such as soybean (Glycine max), rapeseed (including Brassica napus, B. campestris), sunflower (Helianthus annus), cotton (Gossypium hirsutum), corn (Zea mays), cocoa (Theobroma cacao), safflower (Carthamus tinctorius), oil palm (Elaeis guineensis), coconut palm (Cocos nucifera), flax (Linum usitatissimum), and peanut (Arachis hypogaea).

**[0115]** Expression in plants will use regulatory sequences functional in such plants. The expression of foreign genes in plants is well-established (De Blaere et al., Meth. Enzymol. (1987) 153:277-291). The source of the promoter chosen to drive the expression of the fragments of the invention is not critical provided it has sufficient transcriptional activity to accomplish the invention by increasing or decreasing, respectively, the level of translatable mRNA for the glycerolipid desaturases in the desired host tissue. Preferred promoters include (a) strong constitutive plant promoters, such as those directing the 19S and 35S transcripts in cauliflower mosaic virus (Odell et al., Nature (1985) 313:810-812; Hull et al., Virology (1987) 86:482-493), and (b) tissue- or developmentally-specific promoters. Examples of tissue-specific promoters are the light-inducible promoter of the small subunit of ribulose 1,5-bis-phosphate carboxylase (if expression is desired in photosynthetic tissues), the maize zein protein promoter (Matzke et al., EMBO J. (1984) 3:1525-1532), and the chlorophyll a/B binding protein promoter (Lampa et al., Nature (1986) 316:750-752).

**[0116]** Particularly preferred promoters are those that allow seed-specific expression. This may be especially useful since seeds are the primary source of vegetable oils and also since seed-specific expression will avoid any potential deleterious effect in non-seed tissues. Examples of seed-specific promoters include, but are not limited to, the promoters of seed storage proteins, which can represent up to 90% of total seed protein in many plants. The seed storage proteins are strictly regulated, being expressed almost exclusively in seeds in a highly tissue-specific and stage-specific manner (Higgins et al., Ann. Rev. Plant Physiol. (1984) 35:191-221; Goldberg et al., Cell (1989) 56:149-160). Moreover, different seed storage proteins may be expressed at different stages of seed development.

**[0117]** Expression of seed-specific genes has been studied in great detail (See reviews by Goldberg et al., Cell (1989) 56:149-160 and Higgins et al., Ann. Rev. Plant Physiol. (1984) 35:191-221). There are currently numerous examples of seed-specific expression of seed storage protein genes in transgenic dicotyledonous plants. These include genes from dicotyledonous plants for bean b-phaseolin (Sengupta-Gopalan et al., Proc. Natl. Acad. Sci. USA (1985) 82:3320-3324; Hoffman et al., Plant Mol. Biol. (1988) 11:717-729), bean lectin (Voelker et al., EMBO J. (1987) 6: 3571-3577), soybean lectin (Okamuro et al., Proc. Natl. Acad. Sci. USA (1986) 83:8240-8244), soybean Kunitz trypsin inhibitor (Perez-Grau et al., Plant Cell (1989) 1:095-1109), soybean b-conglycinin (Beachy et al., EMBO J. (1985) 4: 3047-3053; pea vicilin (Higgins et al., Plant Mol. Biol. (1988) 11:683-695), pea convicilin (Newbigin et al., Planta (1990) 180:461-470), pea legumin (Shirsat et al., Mol. Gen. Genetics (1989) 215:326-331); rapeseed napin (Radke et al., Theor. Appl. Genet. (1988) 75:685-694) as well as genes from monocotyledonous plants such as for maize 15 kD zein (Hoffman et al., EMBO J. (1987) 6:3213-3221), maize 18 kD oleosin (Lee at al., Proc. Natl. Acad. Sci. USA (1991) 888:6181-6185), barley b-hordein (Marris et al., Plant Mol. Biol. (1988) 10:359-366) and wheat glutenin (Colot et al., EMBO J. (1987) 6:3559-3564). Moreover, promoters of seed-specific genes operably linked to heterologous coding sequences in chimeric gene constructs also maintain their temporal and spatial expression pattern in transgenic plants. Such examples include use of Arabidopsis thaliana 2S seed storage protein gene promoter to express enkephalin peptides in Arabidopsis and B. napus seeds (Vandekerckhove et al., Bio/Technology (1989) 7:929-932), bean lectin and bean b-phaseolin promoters to express luciferase (Riggs et al., Plant Sci. (1989) 63:47-57), and wheat glutenin promoters to express chloramphenicol acetyl transferase (Colot et al., EMBO J. (1987) 6:3559-3564).

**[0118]** Of particular use in the expression of the nucleic acid fragment of the invention will be the heterologous promoters from several soybean seed storage protein genes such as those for the Kunitz trypsin inhibitor (Jofuku et al., Plant Cell (1989) 1:1079-1093; glycinin (Nielson et al., Plant Cell (1989) 1:313-328), and b-conglycinin (Harada et al., Plant Cell (1989) 1:415-425). Promoters of genes for a- and b-subunits of soybean β-conglycinin storage protein will be particularly useful in expressing the mRNA or the antisense RNA in the cotyledons at mid- to late-stages of seed development (Beachy et al., EMBO J. (1985) 4:3047-3053) in transgenic plants. This is because there is very little position effect on their expression in transgenic seeds, and the two promoters show different temporal regulation. The promoter for the a-subunit gene is expressed a few days before that for the b-subunit gene. This is important for transforming rapeseed where oil biosynthesis begins about a week before seed storage protein synthesis (Murphy et

al., J. Plant Physiol. (1989) 135:63-69).

**[0119]** Also of particular use will be promoters of genes expressed during early embryogenesis and oil biosynthesis. The native regulatory sequences, including the native promoters, of the glycerolipid desaturase genes expressing the nucleic acid fragments of the invention can be used following their isolation by those skilled in the art. Heterologous promoters from other genes involved in seed oil biosynthesis, such as those for B. napus isocitrate lyase and malate synthase (Comai et al., Plant Cell (1989) 1:293-300), delta-9 desaturase from safflower (Thompson et al. Proc. Natl. Acad. Sci. USA (1991) 88:2578-2582) and castor (Shanklin et al., Proc. Natl. Acad. Sci. USA (1991) 88:2510-2514), acyl carrier protein (ACP) from Arabidopsis (Post-Beittenmiller et al., Nucl. Acids Res. (1989) 17:1777), B. napus (Safford et al., Eur. J. Biochem. (1988) 174:287-295),and B. campestris (Rose et al., Nucl. Acids Res. (1987) 15:7197), b-ketoacyl-ACP synthetase from barley (Siggaard-Andersen et al., Proc. Natl. Acad. Sci. USA (1991) 88:4114-4118), and oleosin from Zea mays (Lee et al., Proc. Natl. Acad. Sci. USA (1991) 88:6181-6185), soybean (Genbank Accession No: X60773) and B. napus (Lee et al., Plant Physiol. (1991) 96:1395-1397) will be of use. If the sequence of the corresponding genes is not disclosed or their promoter region is not identified, one skilled in the art can use the published sequence to isolate the corresponding gene and a fragment thereof containing the promoter. The partial protein sequences for the relatively-abundant enoyl-ACP reductase and acetyl-CoA carboxylase are also published (Slabas et al.; Biochim. Biophys. Acta (1987) 877:271-280; Cottingham et al., Biochim. Biophys. Acta (1988) 954:201-207) and one skilled in the art can use these sequences to isolate the corresponding seed genes with their promoters. Similarly, the fragments of the present invention encoding glycerolipid desaturases can be used to obtain promoter regions of the corresponding genes for use in expressing chimeric genes.

**[0120]** Attaining the proper level of expression of the nucleic acid fragments of the invention may require the use of different chimeric genes utilizing different promoters. Such chimeric genes can be transferred into host plants either together in a single expression vector or sequentially using more than one vector.

**[0121]** It is envisioned that the introduction of enhancers or enhancer-like elements into the promoter regions of either the native or chimeric nucleic acid fragments of the invention will result in increased expression to accomplish the invention. This would include viral enhancers such as that found in the 35S promoter (Odell et al., Plant Mol. Biol. (1988) 10:263-272), enhancers from the opine genes (Fromm et al., Plant Cell (1989) 1:977-984), or enhancers from any other source that result in increased transcription when placed into a promoter operably linked to the nucleic acid fragment of the invention.

**[0122]** Of particular importance is the DNA sequence element isolated from the gene for the a-subunit of b-conglycinin that can confer 40-fold seed-specific enhancement to a constitutive promoter (Chen et al., Dev. Genet. (1989) 10: 112-122). One skilled in the art can readily isolate this element and insert it within the promoter region of any gene in order to obtain seed-specific enhanced expression with the promoter in transgenic plants. Insertion of such an element in any seed-specific gene that is expressed at different times than the b-conglycinin gene will result in expression in transgenic plants for a longer period during seed development.

**[0123]** The invention can also be accomplished by a variety of other methods to obtain the desired end. In one form, the invention is based on modifying plants to produce increased levels of glycerolipid desaturases by virtue of introducing more than one copy of the foreign gene containing the nucleic acid fragments of the invention. In some cases, the desired level of polyunsaturated fatty acids may require introduction of foreign genes for more than one kind of glycerolipid desaturase.

**[0124]** Any 3' non-coding region capable of providing a polyadenylation signal and other regulatory sequences that may be required for the proper expression of the nucleic acid fragments of the invention can be used to accomplish the invention. This would include 3' ends of the native glycerolipid desaturase(s), viral genes such as from the 35S or the 19S cauliflower mosaic virus transcripts, from the opine synthesis genes, ribulose 1,5-bisphosphate carboxylase, or chlorophyll a/b binding protein. There are numerous examples in the art that teach the usefulness of different 3' non-coding regions.

Transformation Methods

**[0125]** Various methods of transforming cells of higher plants according to the present invention are available to those skilled in the art (see EPO Pub. 0 295 959 A2 and 0 318 341 A1). Such methods include those based on transformation vectors utilizing the Ti and Ri plasmids of Agrobacterium spp. It is particularly preferred to use the binary type of these vectors. Ti-derived vectors transform a wide variety of higher plants, including monocotyledonous and dicotyledonous plants (Sukhapinda et al., Plant Mol. Biol. (1987) 8:209-216; Potrykus, Mol. Gen. Genet. (1985) 199: 183). Other transformation methods are available to those skilled in the art, such as direct uptake of foreign DNA constructs (see EPO Pub. O 295 959 A2), techniques of electroporation (Fromm et al., Nature (1986) (London) 319: 791) or high-velocity ballistic bombardment with metal particles coated with the nucleic acid constructs (Kline et al., Nature (1987) (London) 327:70). Once transformed, the cells can be regenerated by those skilled in the art.

**[0126]** Of particular relevance are the recently described methods to transform foreign genes into commercially im-

portant crops, such as rapeseed (De Block et al., Plant Physiol. (1989) 91:694-701), sunflower (Everett et al., Bio/Technology (1987) 5:1201), and soybean (Christou et al., Proc. Natl. Acad. Sci USA (1989) 86:7500-7504.

Application to RFLP Technology

**[0127]** The use of restriction fragment length polymorphism (RFLP) markers in plant breeding has been well-documented in the art (Tanksley et al., Bio/Technology (1989) 7:257-264). The nucleic acid fragments of the invention can be used as RFLP markers for traits linked to expression of glycerolipid desaturases. These traits will include altered levels of unsaturated fatty acids. The nucleic acid fragment of the invention can also be used to isolate the glycerolipid desaturase gene from variant (including mutant) plants with altered levels of unsaturated fatty acids. Sequencing of these genes will reveal nucleotide differences from the normal gene that cause the variation. Short oligonucleotides designed around these differences may be used as hybridization probes to follow the variation in polyunsaturates. Oligonucleotides based on differences that are linked to the variation may be used as molecular markers in breeding these variant oil traits.

EXAMPLES

**[0128]** The present invention is further defined in the following Examples, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions. All publications, including patents and non-patent literature, referred to in this specification are expressly incorporated by reference herein.

EXAMPLE 1

ISOLATION OF GENOMIC DNA FLANKING THE T-DNA SITE OF INSERTION IN ARABIDOPSIS THALIANA MUTANT LINE 3707

Identification of an Arabidopsis thaliana T-DNA Mutant with Low Linolenic Acid Content

**[0129]** A population of Arabidopsis thaliana (geographic race Wassilewskija) transformants containing the T-DNA of Agrobacterium tumefaciens was generated by seed transformation as described by Feldmann et al., (Mol. Gen. Genetics (1987) 208:1-9). In this population the transformants contain DNA sequences encoding the pBR322 bacterial vector, nopaline synthase, neomycin phosphotransferase (NPTII, confers kanamycin resistance), and b-lactamase (confers ampicillin resistance) within the T-DNA border sequences. The integration of the T-DNA into different areas of the chromosomes of individual transformants may cause a disruption of plant gene function at or near the site of insertion, and phenotypes associated with this loss of gene function can be analyzed by screening the population for the phenotype.

**[0130]** T3 seed was generated from the wild type seed treated with Agrobacterium tumefaciens by two rounds of self-fertilization as described by Feldmann et al., (Science (1989) 243:1351-1354). These progeny were segregating for the T-DNA insertion, and thus for any mutation resulting from the insertion. Approximately 100 seeds of each of 6000 lines were combined and the fatty acid content of each of the 6000 pooled samples was determined by gas chromatography of the fatty acyl methyl esters essentially as described by Browse et al., (Anal. Biochem. (1986) 152: 141-145) except that 2.5% $H_2SO_4$ in methanol was used as the methylation reagent and samples were heated for 1.5 h at 80°C to effect the methanolysis of the seed triglycerides. A line designated "3707" produced seeds that gave an altered fatty acid profile compared to that of the total population. T3 plants were grown from individual T3 seeds produced by line 3707 and self-fertilized to produce T4 seeds on individual plants that were either homozygous wild type, homozygous mutant, or heterozygous for the mutation. The percent fatty acid compositions of a representative subsample of the entire population, of the pooled 3707 T3 seeds, and of a homozygous T4 mutant segregant are shown in Table 4.

TABLE 4

| Fatty Acid Methyl Ester | T3 Pools from lines 3501-4000 average and (std. deviation) | 3707 T3 Pool | 3707 Homozygous T4 Segregant |
|---|---|---|---|
| palmitic | 7.4 (0.37) | 7.0 | 6.4 |

TABLE 4 (continued)

| Fatty Acid Methyl Ester | T3 Pools from lines 3501-4000 average and (std. deviation) | 3707 T3 Pool | 3707 Homozygous T4 Segregant |
|---|---|---|---|
| stearic | 3.0 (0.22) | 2.9 | 3.0 |
| oleic | 17.0 (1.5) | 17.7 | 15.9 |
| linoleic | 29.3 (0.78) | 35.0 | 42.4 |
| linolenic | 16.1 (1.1) | 10.2 | 3.1 |
| eicosenoic | 20.2 (0.73) | 20.5 | 23.6 |

The phenotype of the segregating T3 pool of line 3707 (high linoleic acid, low linolenic acid) was intermediate between that of the population subsample and the homozygous T4 mutant seeds suggesting that line 3707 harbored a mutation at a locus which controls the conversion of linoleic to linolenic acid in the seed. Still, it was not apparent whether the mutant phenotype in line 3707 was the result of a T-DNA insertion. Therefore, Applicants checked a segregating T4 population to determine whether the mutant fatty acid phenotype cosegregated with the nopaline synthase activity and kanamycin resistance encoded by the T-DNA insert. A total of 263 T4 plants were grown and assayed for the presence of nopaline in leaf extracts (Errampalli et al., The Plant Cell (1991) 3:149-157). In addition, T5 seeds were collected from each of the T4 plants and samples of 10-50 seeds were taken to determine the seed fatty acid composition and to determine their ability to germinate in the presence of kanamycin (Feldmann, et al., (1989) Science 243:1351-1354). The 263 plants fell into 3 classes as in Table 5.

TABLE 5

| Number of Individuals | Phenotype |
|---|---|
| 63 | T4 plants: little or no nopaline present; T5 seeds: wild type fatty acid composition, all kanamycin sensitive |
| 134 | T4 plants: nopaline present; T5 seeds: heterozygous fatty acid composition similar to 3707 T3 pool, segregating for kanamycin resistance |
| 64 | T4 plants: nopaline present; T5 seeds homozygous mutant fatty acid composition, all kanamycin resistant |

The cosegregation of the fatty acid phenotype with the phenotypes conferred by T-DNA sequences in an approximately 1:2:1 pattern provided strong evidence that the mutation in line 3707 was the result of a T-DNA insertion. Further experiments were then conducted with the intent of using probes containing T-DNA sequences to clone the T-DNA insert and flanking genomic DNA from line 3707.

Preparation of Genomic DNA from Homozygous 3707 Plants

**[0131]** Seeds from a homozygous line derived from Arabidopsis thaliana (geographic race Wassilewskija (WS)) line 3707 were surface sterilized for 5 min at room temperature in a solution of 5.25% sodium hypochlorite (w/v)/0.15% Tween 20 (v/v), then washed several times in sterile distilled water, with a final rinse in 50% ethanol. Immediately following the ethanol wash, the seeds were transferred to sterile filter paper to dry. One to three seeds were then transferred to 250-mL flasks containing 50 mL of sterile Gamborgs B5 media (Gibco, 500-1153EA), pH 6.0. Cultures were incubated at 22°C, 70 $\mu E \cdot m^{-2} \cdot sec^{-1}$ of continuous light for approximately three weeks, after which time the root tissue was harvested, made into 10 g aliquots (wet weight), lyophilized, and stored at -20°C.

**[0132]** Using a variation of the procedure of Shure et al., (Cell (1983) 35:225-233) genomic DNA was isolated from the root tissue. Two aliquots of lyophilized tissue were ground to a fine powder using a mortar and pestle. The ground tissue was added to a flask containing 85 mL of lysis buffer (7 M urea, 0.35 M NaCl, 0.05 M Tris-HCl, pH 8.0, 0.02 M EDTA, 1% Sarkosyl, 5% phenol) and mixed gently with a glass rod to obtain a homogeneous suspension. To this suspension an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) (equilibrated with 10 mM Tris, pH 8, 1 mM EDTA) was added. After the addition of 8.5 mL of 10% SDS the mixture was swirled on a rotating platform for 15 min at room temperature. After centrifugation at 2000xg for 15 min, the upper aqueous phase was removed to a new tube and extracted two more times, as above, but without the addition of SDS. To the final aqueous phase was added 1/20th the volume of 3 M potassium acetate, pH 5.5 and two times the volume of ice cold 100% ethanol. Precipitation of the DNA was facilitated by incubation at -20°C for one hour followed by centrifugation at 12,000xg for 10 min. The resulting pellet was resuspended in 3 mL of 10 mM Tris, pH 8, 1 mM EDTA to which was added 0.95 g of cesium chloride (CsCl and 21.4 µL of 10 mg/mL ethidium bromide (EtBr) per mL of solution. The DNA was then purified by centrifugation to

equilibrium in a CsCl/EtBr density gradient for 16 h at 15°C, 265,000xg. After removal from the gradient, the DNA was extracted with isopropanol saturated with TE buffer (10 mM Tris, pH 8; 1 mM EDTA) and CsCl to remove EtBr and then dialyzed overnight at 4°C against 10 mM Tris, pH 8, 1 mM EDTA to remove CsCl. The DNA was removed from dialysis and the concentration was determined using the Hoechst fluorometric assay in which an aliquot of DNA is added to 3 mL of 1.5 X $10^{-6}$ M bis-benzimide (Hoechst 33258, Siga) in 1X SSC (0.15 M NaCl, 0.015 M sodium citrate), pH 7.0, incubated at room temperature for 5 min, and read on a fluorometer at excitation 360, emission 450, against a known set of DNA standards.

Plasmid Rescue and Analysis

**[0133]** Five micrograms of genomic DNA from the homozygous 3707 mutant, prepared as described above, was digested with 20 units of either Bam HI or Sal I restriction enzyme (Bethesda Research Laboratory) in a 50 µL reaction volume according to the manufacturer's specifications. After digestion the DNA was extracted with buffer-saturated phenol (Bethesda Research Laboratory) followed by precipitation in ethanol. The resulting pellet was resuspended in a final volume of 10 µL of 10 mM Tris, pH 8, and the concentration of the DNA was determined using the Hoechst fluorometric assay as above.

**[0134]** To facilitate circularization, as opposed to end-to-end joining, a dilute ligation reaction was set up containing 250 ng of Bam HI or Sal I digested genomic DNA, 3 Weiss units of T4 DNA ligase (Promega), 50 µL of 10X ligase buffer (30 mM Tris-HCl, pH 7.8, 100 mM $MgCl_2$, 100 mM DTT, 5 mM ATP) and 5 µL of 100 mM ATP in a 500 µL reaction volume. The reaction was incubated for 16 h at 16°C, heated for 10 min at 70°C, and extracted once with buffer saturated phenol (Bethesda Research Laboratory). The DNA was then precipitated with the addition of two volumes of 100% ethanol and 1/10th volume of 7.5 M ammonium acetate. The resulting pellet was resuspended in a final volume of 10 µL of 10 mM Tris, pH 8, and the concentration of the DNA was determined using the Hoechst fluorometric assay as above.

**[0135]** Competent DH10B cells (Bethesda Research Laboratory) were transfected with 50 ng of ligated DNA at a concentration of 10 ng of DNA per 100 µL of cells according to the manufacturer's specifications. Transformants from Sal I or Bam HI digests were selected on LB plates (10 g Bacto-tryptone, 5 g Bacto-yeast extract, 5 g NaCl, 15 g agar per liter, pH 7.4) containing 100 µg/mL ampicillin or 25 µg/mL kanamycin sulfate, respectively. Ampicillin-resistant ($Amp^r$; ampicillin sensitivity, $Amp^s$) Sal I tranformants were screened for the presence of the kanamycin resistance ($Kan^r$; kanamycin sensitivity, $Kan^s$) gene by picking primary tranformants and stabbing them first to LB plates containing 100 µg/mL ampicillin then to LB plates containing 25 µg/mL kanamycin. After overnight incubation at 37°C the plates were scored for $Amp^r/Kan^s$ colonies. Kanamycin-resistant Bam HI transformants were screened for the presence of the ampicillin resistance gene by picking primary transformants and stabbing them first to LB plates containing 25 µg/mL kanamycin and then to LB plates containing 100 µg/mL ampicillin. After overnight incubation at 37°C the plates were scored for $Kan^r/Amp^r$ colonies.

**[0136]** Cultures were made of 192 $Amp^r/Kan^s$ Sal I transformants and 85 $Kan^r/Amp^r$ Bam HI transformants directly into deep-well microtiter plates containing 200 µL of LB broth (10 g Bacto-tryptone, 5 g Bacto-yeast extract, 5 g NaCl per liter) with 100 µg/mL ampicillin. Using the Schleicher and Schuell Minifold I apparatus and Nytran membranes, dot blots were set up, in duplicate, using the following conditions: 50 µL of culture was diluted into 150 µL of 5X SSC, the culture was lysed and the DNA denatured by the addition of 150 µL of 0.5 M NaOH, 1.5 M NaCl solution for 3 min at room temperature, the filter was removed from the apparatus and neutralized in 0.5 M Tris, pH 8, 1.5 M NaCl, the DNA was then UV cross-linked to the filters using the Stratagene Stratalinker, and the filters were heated for 2 h at 80°C and stored at room temperature.

**[0137]** To determine whether T-DNA was contained within any of the rescued plasmids, the dot blots were probed with portions of the right and left borders of T-DNA. The right border probe consisted of a 2.2 kb Hind III-Dra I fragment of DNA obtained from plasmid H23pKC7 (composed of the 3.2 kb Hind III 23 fragment from Ti plasmid pTiC58 (Lemmers et al., J. Mol. Biol. (1989) 144;353-376) cloned into plasmid vector pKC7 (Maniatis et al., Molecular Cloning, A Laboratory Manual (1982) Cold Spring Harbor Laboratory Press)), and the left border probe consisted of a 2.9 kb Hind III-Eco RI fragment obtained from plasmid H10pKC7 (composed of the 6.5 kb Hind III 10 fragment from Ti plasmid pTiC58 (Lemmers et al., J. Mol. Biol. (1989) 144:353-376) cloned into plasmid vector pKC7 (Maniatis et al., Molecular Cloning, A Laboratory Manual (1982) Cold Spring Harbor Laboratory Press)) using standard digestion, electrophoresis, and electroelution conditions as described in Sambrook et al., (Molecular Cloning, A Laboratory Manual, 2nd ed (1989) Cold Spring Harbor Laboratory Press). Final DNA purification was obtained by passage of the eluted DNA over an Elutip-D column (Schleicher and Schuell) using the manufacturer's specifications. Concentration of the DNA was determined using the Hoechst fluorometric assay as above. Approximately 100 ng of each probe was labeled with a[$^{32}$P] dCTP using a Random Priming Kit from Bethesda Research Laboratories under conditions recommended by the manufacturer. Labeled probe was separated from unincorporated a[$^{32}$P]dCTP by passing the reaction through a Sephadex G-25 spun column under standard conditions as described in Sambrook et al., (Molecular Cloning, A Laboratory Man-

ual, 2nd ed. (1989) Cold Spring Harbor Laboratory Press).

**[0138]** The filters were pre-hybridized in 150 mL of buffer consisting of 6X SSC, 10X Denhardt's solution, 1% SDS, and 100 µg/mL denatured calf thymus DNA for 16 h at 42°C. The denatured, purified, labeled probe was added to the pre-hybridized filters following transfer of the filters to 50 mL of hybridization buffer consisting of 6X SSC, 1% SDS, 10% dextran sulfate, and 50 µg/mL denatured calf thymus DNA. Following incubation of the filters in the presence of the probe for 16 h at 65°C, the filters were washed twice in 150 mL of 6X SSC, 0.5% SDS, twice in 1X SSC, 1% SDS and once in 0.1X SSC, 1% SDS, all at 65°C. The washed filters were subjected to autoradiography on Kodak XAR-2 film at 80°C overnight.

**[0139]** Of the 85 Bam HI candidates, 63 hybridized with the left border probe and none hybridized with the right border probe. Of the 192 Sal I candidates, 31 hybridized with the left border probe, 4 hybridized with the right border probe, and none hybridized with both probes. Twelve of the Bam HI candidates, 7 positive and 5 negative for the presence of the left border of T-DNA, were further analyzed by restriction digests.

**[0140]** DNA from the Bam HI candidates was made by the alkaline lysis miniprep procedure of Birmbiom et al., (Nuc. Acid Res. (1979) 7:1513-1523), as described in Sambrook et al., (Molecular Cloning, A Laboratory Manual, 2nd ed. (1989), Cold Spring Harbor Laboratory Press). The plasmid DNA was digested with Eco RI restriction enzyme (Bethesda Research Laboratories) in accordance with the manufacturer's specifications and electrophoresed through a 0.8% agarose gel in 1X TBE buffer (0.089 M Tris-borate, 0.089 M boric acid, 0.002 M EDTA). All of the Bam HI candidates which hybridized with the left border probe of T-DNA had the same Eco RI restriction pattern, which indicated the presence of 14.2 kb of T-DNA and 1.4 kb of putative plant genomic DNA in these clones.

**[0141]** DNA from Sal I candidates was isolated, restriction-analyzed using Eco RI, Bam HI and Sal I enzymes, and electrophoresed through a 0.8% agarose gel, as above. All of the Sal I candidates which hybridized with the left border probe of T-DNA included 2.9 kb of putative plant DNA. Contained within this 2.9 kb fragment was a 1.4 kb Bam HI-Eco RI fragment as seen with the Bam HI rescued plasmids, suggesting that the 1.4 kb fragment was a subset of the 2.9 kb fragment and that it was adjacent to the left border of the T-DNA at its site of insertion into the plant genome. Sequence analysis of one Sal I candidate (pS1) using a primer homologous to the left border sequence of T-DNA, revealed that the sequence of pS1 was colinear with the sequence of the T-DNA left border (Yadav et al., Proc. Natl. Acad. Sci. USA (1982) 79:6322-6326) up to nucleotide 65, followed by non-T-DNA (putative plant) sequences.

Southern Analysis with Putative Plant DNA from Rescued Plasmids

**[0142]** DNA from the seven Bam HI candidates which hybridized with the left border of the T-DNA was pooled and a portion was digested with Eco RI and Bam HI restriction endonucleases and electrophetically separated on a 0.8% agarose gel in 1X TBE buffer. After excising a 1.4 kb Eco RI-Bam HI fragment from the agarose gel, the 1.4 kb fragment was purified by use of a Gene Clean Kit from Bio 101. Fifty nanograms of the resulting DNA fragment was labeled with a[$^{32}$P]dCTP using a Random Priming Kit (Bethesda Research Laboratory) under conditions recommended by the manufacturer.

**[0143]** Three micrograms of total genomic DNA from homozygous wild-type Arabidopsis and homozygous 3707 mutant Arabidopsis plants was digested to completion with one of the following restriction enzymes: Sal I, Hind III, Eco RI, Cla I, and Bam HI under conditions suggested by the manufacturer. The digested DNA was subjected to electrophoresis and Southern transfer to Hybond-N membranes (Amersham) as described in Sambrook et al. (Molecular Cloning, A Laboratory Approach, 2nd. ed. (1989) Cold Spring Harbor Laboratory Press). After Southern transfer, the membranes were exposed to UV light using the Stratalinker (Stratagene) as per the manufacturer's instructions, air dried, and heated at 68°C for 2 h.

**[0144]** The filters were prehybridized in 1 M NaCl, 50 mM Tris-Cl, pH 7.5, 1% sodium dodecyl sulfate, 5% dextran sulfate, 100 µg/mL of denatured salmon sperm DNA at 65°C overnight. Fifty nanograms of the radiolabeled 1.4 kb Eco RI-Bam HI plant DNA fragment prepared above was added to the prehybridization solution containing the Southern blot and further incubated at 65°C overnight. The filter was washed for 10 min twice in 200 mL 2X SSPE, 0.1% sodium dodecyl sulfate at 65°C and for 10 min in 200 mL 0.5% SSPE, 0.1% sodium dodecyl sulfate at 65°C. Hybridizing fragments were detected by autoradiography. The analysis confirmed that the probe fragment contained plant DNA and that the T-DNA integration site was in a 2.8 kb Bam HI, a 5.2 kb Hind III, a 3.5 kb Sal I, a 5.5 kb Eco RI, and an approximately 9 kb Cla I fragment of wild type Arabidopsis DNA.

Isolation of Lambda Clones Containing the Wild Type Arabidopsis Delta-15 Desaturase Gene

**[0145]** The 1.4 kb Eco RI-Bam HI fragment (see above) was used as a probe to screen a lGem-11 library made from genomic DNA isolated from wildtype Arabidopsis thaliana plants, geographic race WS. To construct the library, genomic DNA was partially digested with Sau3A enzyme, and size-fractionated over a salt gradient as described in Sambrook et al. (Molecular Cloning, A Laboratory Approach, 2nd ed. (1989) Cold Spring Harbor Laboratory Press). The size-

fractionated DNA was then cloned into Bam HI-digested lGem-11 phage DNA (Promega) following the protocol outlined by the manufacturer. About 25,000 plaque-forming units of phage each were plated on five 150 mm petri plates containing a lawn of KW251 cells on NZY agar media (5 g NaCl, 2 g $MgSO_4 \cdot 7H_2O$, 5 g yeast extract, 10 g NZ Amine (casein hydrolysate from ICN Pharmaceuticals), 15 g agar per liter; pH 7.5). The plaques were adsorbed onto nylon membranes (Colony/Plaque Screen, New England Nuclear), in duplicate, and prepared according to the manufacturer's instructions with the addition of a 2 h incubation at 80°C after air drying the filters. The filters were prehybridized at 55°C in hybridization buffer (1% BSA, 0.5 M NaPi, pH 7.2, ($NaH_2PO_4$ and $Na_2HPO_4$), 10 mM EDTA, and 7% SDS) for 4 h, after which time they were transferred to fresh buffer containing the denatured radiolabeled probe (see above) and incubated overnight at 65°C. The filters were rinsed twice with 0.1X SSC, 1% SDS at 65°C for 30 min each and subjected to autoradiography on Kodak XA-R film at 80°C overnight. Seven positively-hybridizing plaques were subjected to plaque purification as described in Sambrook et al., (Molecular Cloning, A Laboratory Manual, 2nd ed. (1989), Cold Spring Harbor Laboratory Press).

**[0146]** Small scale (5 mL) liquid lysates from each of the 7 clones were prepared and titered on KW251 bacteria as described in Sambrook et al. (Molecular Cloning, A Laboratory Manual, 2nd ed (1989), Cold Spring Harbor Laboratory Press). Phage DNA was isolated using a variation of the method of Chisholm (Biotechniques (1989) 7:21-23) in which the initial lysate was made according to Sambrook et al. (Molecular Cloning, A Laboratory Manual, 2nd ed (1989), Cold Spring Harbor Laboratory Press) the concentration of DNase I and RNase I (Sigma) was reduced by half, and the PEG precipitation step was increased to 16 h. Based on restriction analysis using Hind III, Sal I and Xho I enzymes, the original 7 positive phage fell into 5 different classes. While the average insert size was approximately 15 kb, taken together the clones spanned a 40 kb region of genomic DNA. Through restriction mapping using 4 different enzymes (Hind III, Bam HI, Kpn I, and Sal I) singly, and in pair-wise combinations, accompanied by Southern analysis with the 1.4 kb Eco RI-Bam HI probe (as above) and other probes obtained from the 1 clones themselves, a partial map was obtained in which all 5 clones (11111, 141A1, 14211, 14311 and 14411) were found to share an approximately 3 kb region of homology near the site of T-DNA insertion. Via restriction and Southern analysis, Applicants ascertained that a 5.2 kb Hind III fragment present in clones 1111, 41A1, and 4411 also spanned the site of the T-DNA insertion. This fragment was excised from lambda clone 41A1, inserted into the Hind III site of the pBluescript vector (Stratagene), and the resulting plasmid, designated pF1, was prepared and isolated using standard protocols. This Hind III fragment was subsequently used to probe an Arabidopsis cDNA library (see below).

## EXAMPLE 2

### CLONING OF ARABIDOPSIS THALIANA DELTA-15 DESATURASE cDNA USING GENOMIC DNA FLANKING THE T-DNA SITE OF INSERTION IN ARABIDOPSIS THAITANA MUTANT LINE 3707 AS A HYBRIDIZATION PROBE

**[0147]** The 5.2 kb Hind III fragment from plasmid pF1 was purified by electrophoresis in agarose after digestion of the plasmid with Hind III and radiolabeled with $^{32}P$ as described above. For the preparation of an Arabidopsis cDNA library, polyadenylated mRNA was prepared from 3 day-old, etiolated Arabidopsis (ecotype Columbia) seedling hypocotyls using standard protocols (Sambrook, et al., Molecular Cloning: A Laboratory Manual, 2nd Ed. (1989) Cold Spring Harbor Laboratory Press). Five micrograms of this mRNA were used as template with an oligo d(T) primer, and Moloney Murine Leukemia Virus reverse transcriptase (Pharmacia) was used to catalyze first strand cDNA synthesis. Second-strand cDNA was made according to Gubler et al., (Gene (1983) 25:263-272) except that DNA ligase was omitted. After the second strand synthesis, the ends of the cDNA were made blunt by reaction with the Klenow fragment of DNA polymerase and ligated to Eco RI/Not I adaptors (Pharmacia). The cDNA's were purified by spun-column chromatography using Sephacryl S-300 and size-fractionated on a 1% low melting point agarose gel. Size-selected cDNAs (1-3 kb) were removed from the gel using agarase (New England Biolabs) and purified by phenol:chloroform extraction and ethanol precipitation. One hundred nanograms of the cDNA was co-precipitated with 1 µg of 1 ZAP II (Stratagene) Eco RI-digested, dephosphorylated arms. The DNAs were ligated in a volume of 4 µL overnight, and the ligation mix was packaged in vitro using the Gigapack II Gold packaging extract (Stratagene).

**[0148]** Approximately 80,000 phage were screened for positively hybridizing plaques using the radiolabeled 5.2 kb Hind III fragment as a probe essentially as described above and in Sambrook et al., (Molecular Cloning: A Laboratory Manual, 2nd ed. (1989) Cold Spring Harbor Laboratory Press). Replica filters of the phage plaques were soaked in 1 M NaCl, 50 mM Tris-HCl, pH 7.5, 1% SDS, 5% dextran sulfate, 0.1 mg/mL denatured salmon sperm DNA during the pre-hybridization step (8 hr at 65°C) and then probe was added and the hybridization proceeded over 16 hr at the same temperature. Filters were washed sequentially with 2X SSPE, 0.1% SDS at room temperature for 5 min and then again with fresh solution for 10 min, and finally with 0.5X SSPE, 0.1% SDS at 65°C for 5 min. Approximately 20 positively hybridizing plaques were identified in the primary screen. Four of these were picked and subjected to two further rounds of screening and purification. From the tertiary screen, four pure phage plaques were isolated. Plasmid clones containing the cDNA inserts were obtained through the use of a helper phage according to the in vivo excision protocol

provided by Stratagene. Double-stranded DNA was prepared using the alkaline lysis method as previously described, and the resulting plasmids were size-analyzed by electrophoresis in agarose gels. The largest one of these, designated pCF3, contained an approximately 1.4 kb insert which was sequenced using Sequenase T7 DNA polymerase (US Biochemical Corp.) and the manufacturer's instructions, beginning with primers homologous to vector sequences that flank the cDNA insert and continuing serially with primers designed from the newly acquired sequences as the sequencing experiment progressed. The sequence of this insert is shown in SEQ ID NO:1.

EXAMPLE 3

CLONING OF AN ARABIDOPSIS CDNA ENCODING A PLASTID DELTA-15 FATTY ACID DESATURASE

**[0149]** A related fatty acid desaturase was cloned in a similar fashion, except that the probe used was not derived from a PCR reaction on pCF3, but rather was the actual 1.4 kb Not I fragment isolated from pCF3 which was purified and radiolabeled as described above.

**[0150]** Approximately 80,000 phage from the Arabidopsis etiolated hypocotyl cDNA library described above were plated out and screened essentially as before, except as indicated below. The filters were soaked in 1 M NaCl, 50 mM Tris-HCl, pH 7.5, 1% SDS, 5% dextran sulfate, 0.1 mg/mL denatured salmon sperm DNA during the pre-hybridization step (8 hr at 50°C). Then probe was added and the hybridization proceeded over 16 hr at the same temperature. Filters were washed sequentially with 2X SSPE, 0.1% SDS at room temperature for 5 min and then again with fresh solution for 10 min, and finally with 0.5X SSPE, 0.1% SDS at 50°C for 5 min. Approximately 17 strongly hybridizing and 17 weakly hybridizing plaques were identified in the primary screen. Four of the weakly hybridizing plaques were picked and subjected to one to two further rounds of screening with the radiolabeled probe as above until they were pure. To ensure that these were not delta-15 desaturase clones, they were further analyzed to determine whether they hybridized to a delta-15 desaturase 3' end-specific probe. The probe used was an 18 bp oligonucleotide which is complementary in sequence (i.e., antisense) to nucleotides 1229 - 1246 of SEQ ID NO:1. The probe was radiolabeled with gamma-$^{32}$P ATP using T4 polynucleotide kinase and hybridized to filters containing DNA from the isolated clones in 6X SSC, 5X Denhardt's, 0.1 mg/mL denatured salmon sperm DNA, 1 mM EDTA, 1% SDS at 44°C overnight. The filters were washed twice in 6X SSC, 0.1% SDS for 5 min at room temperature, then in 6X SSC, 0.1% SDS at 44°C for 3-5 min. After autoradiography of the filters, one of the clones failed to show hybridization to this probe. This clone was picked, and a plasmid clone containing the cDNA insert was obtained through the use of a helper phage according to the in vivo excision protocol provided by Stratagene. Double-stranded DNA was prepared using the alkaline lysis method as previously described, and the resulting plasmid was size-analyzed by electrophoresis in agarose gels following either Not I digestion or digestion with both Nco I and Bgl II. The results were consistent with the presence in this plasmid, designated pCM2, of an approximately 1.3 kb cDNA insert which lacked a 0.7 kb Nco I - Bgl II fragment characteristic of the Arabidopsis delta-15 desaturase cDNA of pCF3. (This fragment corresponds to the DNA located between the Nco I site at nucleotides 474-479 and the Bgl II site at nucleotides 1164-1169 in SEQ ID NO:1). The complete nucleotide sequence of pCM2 is shown in SEQ ID NO:4.

EXAMPLE 4

CLONING OF PLANT FATTY ACID DESATURASE cDNAs FROM OTHER SPECIES BY HYBRIDIZATION TECHNIQUES

**[0151]** An approximately 1.4 kb fragment containing the Arabidopsis delta-15 desaturase coding sequence of SEQ ID NO:1 was obtained from plasmid pCF3 through the use of the polymerase chain reaction (PCR). Primers (M13(-20) and T7-17mer primers, 1991 Stratagene Catalogue numbers 300303 and 300302, respectively) flanking the pCF3 insert were used in the PCR which was carried out essentially as described in the instructions provided by the vendor in the Perkin-Elmer/Cetus PCR kit. This fragment was digested with Not I to remove vector sequences, purified by agarose gel electrophoresis, and radiolabeled with $^{32}$P as previously described.

EXAMPLE 5

CLONING OF BRASSICA NAPUS SEED cDNAs ENCODING DELTA-15 FATTY ACID DESATURASES

**[0152]** A cDNA library from developing Brassica napus seeds was constructed using the polyadenylated mRNA fraction contained in a polysomal RNA preparation from developing Brassica napus seeds. Polysomal RNA was isolated following the procedure of Kamalay et al., (Cell (1980) 19:935-946) from seeds 20-21 days after pollination. The polyadenylated mRNA fraction was obtained by affinity chromatography on oligo-dT cellulose (Aviv et al., Proc. Natl. Acad.

Sci. USA (1972) 69:1408-1411). Four micrograms of polyadenylated mRNA were reverse transcribed and used to construct a cDNA library in lambda phage (Uni-ZAP™ XR vector) using the protocol described in the ZAP-cDNA™ Synthesis Kit (1991 Stratagene Catalog, Item # 200400).

**[0153]** For the purpose of cloning the Brassica napus seed cDNAs encoding delta-15 fatty acid desaturases, the Brassica napus seed cDNA library was screened several times using the inserts from the Arabidopsis cDNAs pCF3 and pCM2 as radiolabelled hybridization probes. One of the Brasssica napus cDNAs obtained in these screens was used as hybridization probe in a subsequent screen.

**[0154]** For each screening experiment approximately 300,000 phages were screened under low stringency hybridization conditions. The filter hybridizations were carried out in 50 mM Tris pH 7.6, 6X SSC, 5X Denhardt's, 0.5% SDS, 100 ug denatured calf thymus DNA at 50°C overnight and the p[ost hybridization washes were performed in 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45°C for 30 min, and then repeated twice with 0.2X SSC, 0.5% SDS at 50°C for 30 min.

**[0155]** Using the Arabidopsis cDNA insert of pCM2 as a probe in a low stringency screen five strongly hybridizing phages were identified. These phages were purified and excised according to the protocols described in the ZAP-cDNA™ Synthesis Kit and pBluescript II Phagemid Kit (1991 Stratagene Catalog, Item # 200400 and 212205). One of these, designated pBNSF3-f2, contained a 1.3 kb insert. pBNSF3-f2 insert was sequenced completely on both strands. pBNSF3-f2 nucleotide sequence is shown in SEQ ID NO:6. A comparison of this sequence with that of the Arabidopsis thaliana delta-15 desaturase clone (SEQ ID NO:1) confirmed that pBNSF3-f2 is a Brassica napus cDNA that encodes a seed microsomal delta-15 desaturase.

**[0156]** An additional low stringency screen of the Brassica napus seed cDNA library using the cDNA insert in pCM2 as a probe identified eight strongly-hybridizing phages. These phages were plaque purified and used to excise the phagemids as described above. One of these, designated pBNSFd-8, contained a 0.3kb insert. pBNSFd-8 was sequenced completely on one strand, this sequence had significant divergence from the sequence of pBNSF3-f2. The cDNA insert in pBNSFd-8 was used as a hybridization probe in a high stringency screen of the Brassica napus seed cDNA library. The filter hybridizations were carried out in 50 mM Tris pH 7.6, 6X SSC, 5X Denhardt's, 0.5% SDS, 100 ug denatured calf thymus DNA overnight at 50°C and post hybridization washes were in 6X SSC, 0.5% SDS at room temperature for 15 min, then with 2X SSC, 0.5% SDS at 45°C for 30 min, and then twice with 0.2X SSC, 0.5% SDS at 60°C for 30 min. The high stringency screen resulted in three strongly hybridizing phages that were purified and excised as above. One of the excised plasmids pBNSFd-3 contained a 1.4kb insert that was sequenced completely on both strands. SEQ ID NO:8 shows the nucleotide sequence of pBNSFd-3. A comparison of this sequence with that of the Arabidopsis thaliana delta-15 desaturase clone (SEQ ID NO:4) confirmed that pBNSFd-3 is a Brassica napus cDNA that encodes a seed plastid delta-15 desaturase.

Cloning of a Soybean Seed cDNA Encoding a Microsomal Delta-15 Glycerolipid Desaturase

**[0157]** A cDNA library was made as follows: Soybean embryos (ca. 50 mg fresh weight each) were removed from the pods and frozen in liquid nitrogen. The frozen embryos were ground to a fine powder in the presence of liquid nitrogen and then extracted by Polytron homogenization and fractionated to enrich for total RNA by the method of Chirgwin et al. (Biochemistry (1979) 18:5294-5299). The nucleic acid fraction was enriched for poly $A^+$RNA by passing total RNA through an oligo-dT cellulose column and eluting the poly $A^+$RNA with salt as described by Goodman et al. (Meth. Enzymol. (1979) 68:75-90). cDNA was synthesized from the purified poly $A^+$RNA using cDNA Synthesis System (Bethesda Research Laboratory) and the manufacturer's instructions. The resultant double-stranded DNA was methylated by Eco RI DNA methylase (Promega) prior to filling-in its ends with T4 DNA polymerase (Bethesda Research Laboratory) and blunt-end ligation to phosphorylated Eco RI linkers using T4 DNA ligase (Pharmacia). The double-stranded DNA was digested with Eco RI enzyme, separated from excess linkers by passage through a gel filtration column (Sepharose CL-4B), and ligated to lambda ZAP vector (Stratagene) cording to manufacturer's instructions. Ligated DNA was packaged into phage using the Gigapack packaging extract (Stratagene) according to manufacturer's instructions. The resultant cDNA library was amplified as per Stratagene's instructions and stored at -80°C.

**[0158]** Following the instructions in the Lambda ZAP Cloning Kit Manual (Stratagene), the cDNA phage library was used to infect E. coli BB4 cells and approximately 80,000 plaque forming units were plated onto 150 mm diameter petri plates. Duplicate lifts of the plates were made onto nitrocellulose filters (Schleicher & Schuell). The filters were prehybridized in 25 mL of hybridization buffer consisting of 50mM Tris-HCl, pH 7.5, 1 M NaCl, 1% SDS, 5% dextran sulfate and 0.1 mg/mL denatured salmon sperm DNA (Sigma Chemical Co.) at 50°C for 2 h. Radiolabeled probe prepared from pCF3 as described above was added, and allowed to hybridize for 18 h at 50°C. The probes were washed twice at room temperature with 2X SSPE, 1% SDS for five minutes followed by washing for 5 min at 50°C in 0.2X SSPE, 1% SDS. Autoradiography of the filters indicated that there was one strongly hybridizing plaque, and approximately five weakly hybridizing plaques. The more strongly hybridizing plaque was subjected to a second round of screening as before, excepting that the final wash was for 5 min at 60°C in 0.2X SSPE, 1% SDS. Numerous, strongly hybridizing

plaques were observed, and one, well-isolated from other phage, was picked for further analysis.

**[0159]** Following the Lambda ZAP Cloning Kit Instruction Manual (Stratagene), sequences of the pBluescript vector, including the cDNA inserts, from the purified phage was excised in the presence of a helper phage and the resultant phagemid was used to infect E. coil XL-1 Blue cells. DNA from the plasmid, designated pXF1, was made by the alkaline lysis miniprep procedure described in Sambrook et al. (Molecular Cloning, A Laboratory Manual, 2nd ed. (1989) Cold Spring Harbor Laboratory Press). The alkali-denatured double-stranded DNA from pXF1 was completely sequenced on both strands. The insert of pXF1 contained a stretch of 1783 nucleotides which contained an unknown open-reading frame and also contained a poly-A stretch of 16 nucleotides 3' to the open reading frame, from nucleotides 1767 to 1783, followed by an Eco RI restriction site. The 2184 bases that followed this Eco RI site contained a 1145 bp open reading frame which encoded a polypeptide of about 68% identity to, and colinear with, the Arabidopsis delta-15 desaturase polypeptide listed in SEQ ID No:2. The putative start methionine of the 1145 bp open-reading frame corresponded to the start methionine of the Arabidopsis microsomal delta-15 peptide and there were no amino acids corresponding to a plastid transit peptide 5' to this methionine. When the insert in pXF1 was digested with Eco RI four fragments were observed, fragments of approximately 370 bp and 1400 bp fragments, derived from the first 1783 bp of the insert in pXF1, and fragments of approximately 600 bp and 1600 bp derived from the the other 2184 nucleotides of the insert in pXF1. Only the 600 bp and 1600 bp fragments hybridized with probe derived from pCF3 on Southern blots. It was deduced that pXF1 contained two different cDNA inserts separated by an Eco RI site and the second of these inserts was a 2184 bp cDNA encoding a soybean microsomal delta-15 desaturase. The complete nucleotide sequence of the 2184 bp soybean microsomal delta-15 cDNA contained in plasmid pXF1 is listed in SEQ ID No:10.

Cloning of a Soybean Seed cDNA Encoding a Plastid Delta-15 Glycerolipid Desaturase Using Soybean Microsomal Delta-15 Desaturase cDNA as an Hybridization Probe

**[0160]** A 1.0 kb fragment of the coding region of the soybean microsomal delta-15 desaturase cDNA contained in plasmid pXF1 was excised by digestion with the restriction enzyme Hha I. This 1.0 Kb fragment was purified by agarose gel electrophoresis and radiolabeled with 32P as previously described. The radiolabeled fragment was used to screen 100,000 plaque-forming units of the the soybean cDNA library as described above. Autoradiography of the filters indicated that there were eight hybridizing plaques and these were subjected to a second round of screening. Sequences of the pBluescript vector from all eight of the purified phages, including the cDNA inserts, were excised in the presence of a helper phage and the resultant phagemids were used to infect E. coli XL-1 Blue cells. DNA from the plasmids was made by the alkaline lysis miniprep procedure described in Sambrook et al. (Molecular Cloning, A Laboratory Manual, 2nd ed. (1989) Cold Spring Harbor Laboratory Press). Restriction analysis showed they contained inserts ranging from 1.0 kb to 3.0 kb in size. One of these inserts, designated pSFD-118bwp, contained an insert of about 1700 bp. The alkali-denatured double-stranded DNA from pSFD-118bwp was completely sequenced on both strands. The insert of pSFD-118bwp contained a stretch of 1675 nucleotides which contained an open-reading frame encoding a polypeptide of about 80% identity with, and colinear with, the Arabidopsis plastid delta-15 desaturase polypeptide listed in SEQ ID No:5. The open-reading frame also encoded amino acids corresponding to a plastid transit peptide at the 5' end of the open-reading frame. The transit peptide was colinear with, and shared some homology to, the transit peptide described for the Arabidopsis plastid delta-15 glycerolipid desaturase. Based on the homology to Arabidopsis plastid delta-15 glycerolipid desaturase and because of the presence of a plastid transit peptide, the cDNA contained in plasmid pSFD-118bwp was deduced to be a soybean plastid delta-15 glycerolipid desaturase. The complete nucleotide sequence of the 1675 bp soybean plastid delta-15 glycerolipid desaturase cDNA is listed in SEQ ID NO:12.

## EXAMPLE 6

## CLONING OF cDNA SEQUENCES ENCODING FATTY ACID DESATURASES BY POLYMERASE CHAIN REACTION

**[0161]** Analysis of the deduced protein sequences of the different higher plant glycerolipid desaturases described in this invention reveals to those skilled in the art regions of the amino acid sequences that have been conserved among higher plants and between higher plants and cyanobacterial des A. These short stretches of amino acids can be used to design oligomers as primers for polymerase chain reactions. Two amino acid sequences that are highly conserved between the des A and plant delta-15 desaturases polypeptides are amino acid sequences 97-108 and 299-311 (SEQ ID NO:2). Polymerase chain reactions (PCRs) were performed using GeneAmp® RNA PCR Kit (Perkin Elmer Cetus) following manufacturer's protocols. In one PCR experiment, SEQ ID NOS:22 and 23 were used as sense primers and either SEQ ID NOS:24 and 25 or SEQ ID NOS:26 and 27 as antisense primers on poly A+ RNA purified from both Arabidopsis leaf and canola developing seeds. For this, ca. 100 ng of polyA+ RNA was isolated as described previously and reverse-transcribed using the kit using random hexamers. Then the cDNA was used in PCR using 64 pmoles each of SEQ ID NOS:22 and 23 as sense primers and either a mixture of 64 pmoles of SEQ ID NO:24 and 78 pmoles of

SEQ ID NO:25 or a mixture 35 pmoles of SEQ ID NO:26 and 50 pmoles of SEQ ID NO:27 by the following program: a) 1 cycle of 2 min at 95°C and 15 C at 50°C, b) 30 cycles of 3 min at 65°C (extension), 1 min 20 sec at 95°C (denaturation), 2 min at 50°C (annealing), and c) 1 cycle of 7 min at 65°C. PCR products were analyzed by gel electrophoresis. All PCRs resulted in PCR products of the correct size (ca. 630 bp). The PCR products from Arabidopsis and canola were purified and used as radiolabeled hybridization probes to screen the Lambda Yes Arabidopsis cDNA library at low stringency, as described above. This led to the isolation of a pure phage, which was excised to give plasmid pYacp7. The cDNA insert in pYacp7 was partially sequenced. Its sequence showed that it encoded an incomplete desaturase polypeptide that was identical to another cDNA (in plasmid pFadx-2) isolated by low-stringency hybridization as described previously. The composite sequence derived from the partial sequences from the cDNA inserts in pFadx-2 and pYacp7 is shown in SEQ ID NO:16 and the polypeptide encoded by it in SEQ ID NO:17. As discussed previously, SEQ ID NO:17 is a putative plastid delta-15 desaturase. A full-length version of pYacp7 can be readily isolated using it has a hybridization probe.

**[0162]** Two additional conserved regions correspond to aminoacid residues 130 to 137 and 249 and 256 of SEQ ID NO:7 (Brassica napus glycerolipid desaturase delta-15). Degenerate oligomers were designed to these regions with additional nucleotides containing a restriction site for Bam H1 were added to the 5' ends of each oligonucleotide to facilitate subcloning of the PCR products. The nucleotide sequences of these oligonucleotides named F2-3 and F2-3c are shown in SEQ ID NO:18 and SEQ ID NO:19 respectively.

**[0163]** Mixtures of degenerate oligonucleotides F2-3 and F2-3c were used to amplify, isolate and clone glycerolipid desaturase sequences represented in corn seed mRNA population, essentially as described in the GeneAmp RNA PCR Kit purchased from Perkin Elmer Cetus and in Innis, et al., Eds, (1990) PCR Protocols: A Guide to Methods and Applications, Academic Press, San Diego.

**[0164]** Corn seed RNA was obtained from developing corn seeds 15-20 days after pollination by the method of Chirgwin et al., (1979) Biochemistry 18:5294. Corn seed polyadenylated mRNA was isolated by affinity chromatography on oligo-dT cellulose (Aviv et al., Proc. Natl. Acad. Sci. USA (1972) 69:1408-1411). 20-50ng of A+mRNA were used in reverse transcription reactions with oligo-dT and random hexamers primers using the reaction buffer and conditions recomended by Perkin Elmer Cetus. The resulting cDNA was then used as template for the amplification of corn seed glycerolipid sequences using the set of degenerate primers in SEQ ID NO: 18 and 19. Reaction conditions were as described by Perkin Elmer Cetus, the amplification protocol consisted of a sequence of 95°C/1 min, 55°C/1 min, 72°C/2 min for 30-50 cycles. The resulting polymerase reaction products were phenol-chloroform extracted, digested with Bam HI and separated from unincorporated primers by gel filtration chromatopgraphy on Linker 6 spin columns (Pharmacia Inc.). The resulting PCR products were cloned into pBluescript SK at the Bam H1 site, and transformed into E. coli DH5 competent cells. Restriction analysis of plasmid DNA from the transformed colonies obtained revealed a colony, PCR-20, that contained an insert of about 0.5 kB in size at the pBluescript SK BamH1 site. The PCR-20 insert was completely sequenced on both strands. The nucleotide sequence of PCR20 insert is shown in SEQ ID NO:14 and the translated amino acid sequence is shown in SEQ ID NO:15. This aminoacid sequence shows an overall identity of 61.9% to the aminoacid sequence of Brassica napus microsomal delta-15 deaturase shown in SEQ ID NO:7. This result identifies the PCR20 insert as a polymerase reaction product of a corn seed delta-15 desaturase cDNA. PCR20 insert may be used as a probe to readily isolate full length corn seed delta-15 desaturase cDNAs or as such to antisense or cosuppress corn seed glycerolipid delta-15 desaturase gene expression in transgenic corn plants by cloning it in the appropriate corn gene expression vector.

## EXAMPLE 7

## USE OF THE ARABIDOPSIS THALIANA DELTA-15 DESATURASE GENOMIC CLONES AS A RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP) MARKERS TO MAP THE DELTA-15 DESATURASE LOCI IN ARABIDOPSIS

**[0165]** DNA flanking the T-DNA insertion site in mutant line 3707 was used to map the genetic locus encoding the delta-15 desaturase of Arabidopsis thaliana seeds. An approximately 12 kB genomic DNA fragment containing the Arabidopsis delta-15 desaturase coding sequence was removed from the lambda-4211 clone by digestion with restriction endonuclease Xho I, separated from the Lambda arms by agarose gel electrophoresis, and purified using standard procedures. The isolated DNA was labeled with $^{32}P$ using a random priming kit from Pharmacia under conditions recommended by the manufacturer. The radioactive DNA was used to probe a Southern blot containing genomic DNA from Arabidopsis thaliana (ecotype Wassileskija and marker line W100 ecotype Landesberg background) digested with one of several restriction endonucleases. Following hybridization and washes under standard conditions (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed. (1989) Cold Spring Harbor Laboratory Press), autoradiograms were obtained. Different patterns of hybridization (polymorphisms) were identified in digests using restriction endonucleases Bgl II, Cla I, Hind III, Nsi I, and Xba I. The same radiolabeled DNA fragment was used to map the

polymorphism essentially as described by Helentjaris et al., (Theor. Appl. Genet. (1986) 72:761-769). The radiolabeled DNA fragment was applied as described above to Southern blots of Xba I digested genomic DNA isolated from 117 recombinant inbred progeny (derived from single-seed descent lines to the $F_6$ generation) resulting from a cross between Arabidopsis thaliana marker line W100 and ecotype Wassileskija (Burr et al., Genetics (1988) 118:519-526). The bands on the autoradiograms were interpreted as resulting from inheritance of either paternal (ecotype Wassileskija) or maternal (marker line W100) DNA or both (a heterozygote). The resulting segregation data were subjected to genetic analysis using the computer program Mapmaker (Lander et al., Genomics (1987) 1:174-181). In conjunction with previously obtained segregation data for 63 anonymous RFLP markers and 9 morphological markers in Arabidopsis thaliana (Chang et al., Proc. Natl. Acad. Sci. USA (1988) 85:6856-6860; Nam et al., Plant Cell (1989) 1:699-705), a single genetic locus was positioned corresponding to the genomic DNA containing the delta-15 desaturase coding sequence. The location of the delta-15 desaturase gene was thus determined to be on chromosome 2 between the lambda AT283 and cosmid c6842 RFLP markers, near the py and erecta morphological markers.

**[0166]** The cDNA in plasmid pCM2 was also shown to hybridize polymorphically to genomic DNA from Arabidopsis thaliana (ecotype Wassileskija and marker line W100 ecotype Landesberg background) digested with Eco RI. It was used as a RFLP marker to map the genetic locus for the gene encoding this fatty acid desaturase in Arabidopsis as described above. A single genetic locus was positioned corresponding to this desaturase cDNA. Its location was thus determined to be on chromosome 3 between the lambda AT228 and cosmid c3838 RFLP markers, "north" of the glabrous locus (Chang et al., Proc. Natl. Acad. Sci. USA (1988) 85:6856-6860; Nam et al., Plant Cell (1989) 1:699-705).

## EXAMPLE 8

### USE OF SOYBEAN SEED MICROSOMAL DELTA-15 GLYCEROLIPID DESATURASE cDNA SEQUENCE IN PLASMID AS A RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP) MARKER

**[0167]** A 600 bp fragment of the cDNA insert from plasmid pXF1, which contains about 300 bp of the coding sequence and 300 bp of the 3' untranslated sequence, was excised by digestion with restriction enzyme Eco RI in standard conditions as described in Sambrook et al. (Molecular Cloning, A Laboratory Manual, 2nd ed. (1989) Cold Spring Harbor Laboratory Press), purified by agarose gel electrophoresis and labeled with $^{32}P$ using a Random Priming Kit from Bethesda Research Laboratories under conditions recommended by the manufacturer. The resulting radioactive probe was used to probe a Southern blot (Sambrook et al., Molecular Cloning, A Laboratory Manual, 2nd ed. (1989) Cold Spring Harbor Laboratory Press) containing genomic DNA from soybean [Glycine max (cultivar Bonus) and Glycine soja (PI81762)], digested with one of several restriction enzymes. After hybridization and washes under standard conditions (Sambrook et al. Molecular Cloning, A Laboratory Manual, 2nd ed. (1989), Cold Spring Harbor Laboratory Press), autoradiograms were obtained and different patterns of hybridization (polymorphisms) were identified in digests performed with restriction enzymes Bam HI, Eco RV and Eco RI. The same probe was then used to map the polymorphic pXF1 locus on the soybean genome, essentially as described by Helentjaris et al. (Theor. Appl. Genet. (1986) 72: 761-769). Plasmid pXF1/600 bp probe was applied, as described above, to Southern blots of EcoRI, PstI, EcoRV, BamHI, or Hin DIII digested genomic DNAs isolated from 68 F2 progeny plants resulting from a G. max Bonus x G. soja PI81762 cross. The bands on the autoradiograms were interpreted as resulting from the inheritance of either paternal (Bonus) or maternal (PI81762) pattern, or both (a heterozygote). The resulting data were subjected to genetic analysis using the computer program Mapmaker (Lander et al., Genomics (1987) 1:174-181). In conjunction with previously obtained data for 436 anonymous RFLP markers in soybean (Tingey et al., J. Cell. Biochem., Supplement 14E (1990) p. 291, abstract R153], Applicants were able to position a single genetic locus corresponding to the pXF1/600 bp probe on the soybean genetic map. This confirms that the gene for microsomal delta-15 desaturase is located on chromosome 19 in the soybean genome. This information will be useful in soybean breeding targeted towards developing lines with altered polyunsaturate levels.

## EXAMPLE 9

### OVEREXPRESSION OF MICROSOMAL DELTA-15 FATTY ACID DESATURASE IN PLANTS

**[0168]** Detailed procedures for DNA manipulation, such as use of restriction endonucleases and other DNA modifying enzymes, agarose gel electrophoresis, isolation of DNA from agarose gels, transformation of E. coli cells with plasmid DNA, and isolation and sequencing of plasmid DNA are described in Sambrook et al. (1989) Molecular cloning, A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press and Ausubel et al. (1989) Current Protocols in Molecular Biology John Wiley & Sons. All restriction enzymes and modifying enzymes were obtained from Bethesda Research Laboratory, unless otherwise noted.

**[0169]** To test the biological effect of overexpression of the microsomal delta-15 desaturase SEQ ID NO:1, i.e., the

cDNA encoding Arabidopsis thaliana microsomal delta-15 desaturase, was placed in the sense orientation behind either the CaMV 35S promotor, to provide constituitive expression, or behind the promotor for the gene encoding soybean a' subunit of the β-conglycinin (7S) seed storage protein, to provide embryo-specific expression. To create the chimeric gene constructs, specific expression cassettes were made to facilitate easy manipulation of the desired clones. The chimeric genes were then transformed into plant cells by Agrobacterium tumefaciens's binary Ti plasmid vector system (Hoekema et al,. (1983) Nature 303:179-180; Bevan (1984) Nucl. Acids Res. 12:8711-8720].

Overexpression of Arabidopsis Delta-15 Fatty Acid Desaturase in Transgenic Carrot Hairy Roots

**[0170]** To confirm the identity of SEQ ID NO:1 (Arabidopsis microsomal delta-15 fatty acid desaturase) and to test the biological effect of its overexpression in a heterologous plant species, the constitutive chimeric gene 35S:SEQ ID NO:1 was introduced into carrot tissue by Agrobacterium. The cassette for constitutive gene expression in plasmid, pAW28, originated from pK35K which, in turn, is derived from pKNK. Plasmid pKNK is a pBR322-based vector containing a chimeric gene for plant kanamycin resistance: nopaline synthase (NOS) promoter/neomycin phosphotransferase (NPT) II coding region/3' NOS chimeric gene. Plasmid pKNK has been deposited on 7 January 1987 with the American Type Culture Collection of Rockville,Maryland, USA under the provisions of the Budapest Treaty and bears the deposit accession number 67284. A map of this plasmid is shown in Lin, et al., Plant Physiol. (1987) 84:856-861. The NOS promoter region is a 296 bp Sau 3A-Pst I fragment corresponding to nucleotides -263 to +33, with respect to the transcription start site, of the NOS gene described by Depicker et al. (1982) J. Appl. Genet. 1:561-574. The Pst I site at the 3' end was created at the translation initiation codon of the NOS gene. The NptII coding region is a 998 bp Hind III-Bam HI fragment obtained from transposon Tn5 (Beck et al., Gene (1982) 19:327-336) by the creation of Hind III and Bam HI sites at nucleotides 1540 and 2518, respectively. The 3' NOS is a 702 bp Bam HI-Cla I fragment from nucleotides 848 to 1550 of the 3' end of the NOS gene (Depicker et al., J. Appl. Genet. (1982) 1:561-574) including its' polyadenylation region. pKNK was converted to pK35K by replacing its Eco RI-Hind III fragment containing the NOS promoter with a Eco RI-Hind III fragment containing the CaMV 35S promoter. The Eco RI-Hind III 35S promoter fragment is the same as that contained in pUC35K that has been deposited on 7 January 1987 with the American Type Culture Collection under the provisions of the Budapest Treaty and bears the deposit accession number 67285. The 35S promoter fragment was prepared as follows, and as described in Odell et al., Nature (1985) 313:810-813, except that the 3' end of the fragment includes CaMV sequences to +21 with respect to the transcription start site. A 1.15 KB Bgl II segment of the CaMV genome containing the region between -941 and +208 relative to the 35S transcription start site was cloned in the Bam HI site of the plasmid pUC13. This plasmid was linearized at the Sal I site in the polylinker located 3' to the CaMV fragment and the 3' end of the fragment was shortened by digestion with nuclease Bal3l. Following the addition of Hind III linkers, the plasmid DNA was recircularized. From nucleotide sequence analysis of the isolated clones, a 3' deletion fragment was selected with the Hind III linker positioned at +21. The 35S promoter fragment was isolated as an Eco RI-Hind III fragment, the Eco RI site coming from the polylinker of pUC13.

**[0171]** The NPTII coding region in plasmid pK35K was removed from plasmid pK35K by digestion with Hind III and Bam HI restriction enzymes. Following digestion, the ends of the DNA molecules were filled-in using Klenow enzyme. Not I linkers (New England Biolabs) were then ligated on the ends and the plasmid was recircularized to yield plasmid pK35Nt. A 1.7 kB fragment containing the 35S promotor region - Not I site - 3' untranslated region from nopaline synthase was liberated from pK35Nt using restriction endonucleases Eco RI and Cla I. Following restriction digestion the ends of the DNA molecules were filled-in using Klenow enzyme after which Xho I linkers (New England Biolabs) were added. The 1.7 kB fragment, now containing Xho I sites at either end, was gel isolated and cloned into the plasmid vector pURA3 (Clonetech) at its unique Xho I site. The vector pURA3 was choosen due to the absence of a Not I restriction site, the presence of a single Xho I restriction site and because the relatively large size of the vector (pURA3) would make the isolation of the gene expression cassettes relatively easy from the final construct.

**[0172]** The 1.4 kB Not I fragment in plasmid pCF3 containing Arabidopsis microsomal delta-15 desaturase (SEQ ID NO:1) was isolated and ligated to pAW28 (the constituitive expression cassette) previously linearized with Not I restriction enzyme and treated with calf intestinal alkaline phosphatase (Boehringer Mannheim) to result in plasmids pAW29 and pAW30 that had SEQ ID NO:1 cloned in a sense orientation and antisense orientation, respectively, with respect to the promoter. The orientation of the cDNA relative to the promotors was established by digestion with appropriate restriction endonucleases or by sequencing across the promotor-cDNA junctions.

**[0173]** The chimeric genes 35S promotor/sense SEQ ID NO:1/3'NOS and 35S promotor/antisense SEQ ID NO: 1/3'NOS were isolated as a 3 kB Xho I fragment from plasmids pAW29 and pAW30, respectively, and cloned into the binary vector pZS194b at its unique Sal I site to result in plasmids pAW31 and pAW32, respectively. The orientation of the plant selectable marker gene in pAW31 and pAW32 is the same as that of the 35S promoter as acertained by digestion with appropriate restriction endonucleases. Binary vector pZS194b contains the pBR322 origin of replication, the replication and stability regions of the Pseudomonas aeruginosa plasmid pVS1 [Itoh, et al., (1984) Plasmid 11: 206-220] required for replication and maintenance of the plasmid in Agrobacterium, the bacterial NPT II gene (kan-

amycin resistance) from Tn5 [Berg et al., (1975) Proc. Mat'l. Acad. Sci. U.S.A. 72:3628-3632] as a selectable marker for transformed bacteria, left and right borders of the T-DNA of the Ti plasmid [Bevan et al., (1984) Nucl. Acids Res. 12:8711-8720], and, between the left and right T-DNA borders are the chimeric NOS:NPT II gene for plant kanamycin resistance, described above, as a selectable marker for transformed plant cells and the E. coli lacZ a-complementing segment [Vieria and Messing (1982) Gene 19:259-267] with unique restriction endonuclease sites for Kpn I and Sal I.

**[0174]** The binary vectors pAW31 and pAW32 were transformed by the freeze/thaw method [Holsters et al. (1978) Mol. Gen. Genet. 163:181-187] into Agrobacterium tumefaciens strain R1000, carrying the Ri plasmid pRiA4b from Agrobacterium rhizogenes [Moore et al., (1979) Plasmid 2:617-626] to result in transformants R1000/pAW31 and R1000/pAW32, respectively.

**[0175]** Carrot (Daucus carota L.) cells were transformed by co-cultivation of carrot root disks with strain R1000, R1000/pAW31, or R1000/pAW32 by the method of Petit et al., (1986) Mol. Gen. Genet. 202:388-393]. To prepare explants for inoculation, carrots purchased from the local supermarket were first scrubbed gently with water and dish detergent, then rinsed thoroughly with tap and distilled water. They were surface sterilized in a stirred solution of 50% Clorox and distilled water for 30 min and rinsed thoroughly with sterile distilled water. The carrots were peeled using an autoclaved vegetable peeler and then sliced with a scalpel blade into disks of approximately 5-10 mm thickness. The disks were placed in petri dishes, onto a medium consisting of distilled deionized water solidified with 0.7% agar, in an inverted orientation so that the cut surface nearest to the root apex of the carrot was exposed for inoculation.

**[0176]** Cultures of Agrobacterium strains R1000, R1000/pAW31, and R1000/pAW32 were initiated from freshly grown plates in LB broth plus the appropriate antibiotic selective agents (50 mg/L chloramphenicol for the R1000 or 50 mg/L each of chloramphenicol and kanamycin for R1000/pAW31 and R1000/pAW32) and grown at 28°C to an optical density of around 1.0 at 600 nm. Bacterial cells were pelleted by centrifugation, rinsed and resuspended in LB broth without antibiotics. Freshly cut carrot disks were inoculated by applying 100 µL of the bacterial suspension to the cut surface of each disk. As a control, some disks were inoculated with sterile LB broth only, to indicate the extent of root formation in the absence of Agrobacterium.

**[0177]** Inoculated root disks were incubated at 25°C in the dark in petri dishes sealed with Parafilm. After two weeks of co-cultivation of carrot disks with Agrobacterium, the carrot disks were transferred to fresh agar-solidified water medium containing 500 mg/L carbenicillin for the counterselection of Agrobacterium. At this time, hairy root formation was noted on some root disks. Transfer of the explants to fresh counterselection medium was done at four weeks. Excision of individual roots from the explants was begun at six weeks. Ten days later, additional roots were taken from the explants as needed.

**[0178]** Approximately 5-10 mm long hairy roots were excised and individually subcultured on MS minimal organics medium with 30 g/L sucrose (Gibco, Grand Island, N. Y., Cat. No. 510-1118EA) and 500 mg/L carbenicillin. Approximately equal numbers of roots were subcultured in liquid medium and in a medium solidified with 0.6% agarose. Cultures on solid medium were grown in 60 x 100 mm petri dishes, liquid cultures were in 6-well culture dishes. When excising roots, an effort was made to select single roots from distinct callus-like outgrowths on the wounded surface. These sites of excision were marked on the lid of the petri dish to minimize repeat sampling of tissue originating from the same transformation event.

**[0179]** Two to three weeks after excision from the explants, individual hairy root cultures that were not visibly contaminated with Agrobacterium were transferred to fresh MS medium supplemented with 500 mg/L carbenicillin. The root mass of each culture was cut into segments including one or more branch roots, and these segments were transferred as a group to a plate or well of fresh medium. Approximately 20 mg fresh weight of tissue of root cultures which grew to adequate size within the next two to three weeks were sampled for fatty acid composition by gas chromatography of the fatty acyl methyl esters essentially as described by Browse et al., (Anal. Biochem. (1986) 152:141-145) except that 2.5% $H_2SO_4$ in methanol was used as the methylation reagent and samples were heated for 1.5 h at 80°C to effect the methanolysis of the seed triglycerides. The results are shown in Table 6. A second sample of tissue consisting of an actively growing root tip of approximately 1 cm was excised and placed on MS medium supplemented with 500 mg/L carbenicillin and 25-50 mg/L kanamycin to test for kanamycin resistance select for hairy roots co-transformed with the binary vector [Simpson et al. (1986) Plant Mol. Biol. 6:403-415].

TABLE 6

| Percent 18:3 and 18:2/18:3 Ratio in Roots of Transgenic Carrots | | | |
|---|---|---|---|
| Root Sample | Transformation Vector Used | %18:3 | %18:2/18:3 |
| 1 | R1000/pAW31 | 62 | 0.09 |
| 2 | R1000/pAW31 | 8 | 7.30 |
| 3 | R1000/pAW31 | 10 | 5.69 |
| 4 | R1000/pAW31 | 62 | 0.06 |

TABLE 6 (continued)

| Percent 18:3 and 18:2/18:3 Ratio in Roots of Transgenic Carrots | | | |
|---|---|---|---|
| Root Sample | Transformation Vector Used | %18:3 | %18:2/18:3 |
| 5 | R1000/pAW31 | 10 | 5.07 |
| 6 | R1000/pAW31 | 4 | 14.2 |
| 7 | R1000/pAW31 | 61 | 0.18 |
| 8 | R1000/pAW31 | 4 | 15.1 |
| 9 | R1000/pAW31 | 61 | 0.07 |
| 10 | R1000/pAW31 | 63 | 0.09 |
| 11 | R1000/pAW31 | 15 | 3.04 |
| 12 | R1000/pAW31 | 64 | 0.14 |
| 13 | R1000/pAW31 | 5 | 9.94 |
| 14 | R1000/pAW31 | 9 | 6.72 |
| 15 | R1000/pAW31 | 8 | 7.08 |
| 16 | R1000/pAW31 | 8 | 6.31 |
| 17 | R1000/pAW31 | 23 | 1.86 |
| 18 | R1000/pAW31 | 8 | 7.33 |
| 19 | R1000/pAW31 | 10 | 5.99 |
| 20 | R1000/pAW31 | 7 | 8.83 |
| 21 | R1000/pAW32 | 9 | 6.80 |
| 22 | R1000/pAW32 | 4 | 11.8 |
| 23 | R1000/pAW32 | 3 | 18.8 |
| 24 | R1000/pAW32 | 10 | 6.21 |
| 25 | R1000/pAW32 | 7 | 8.57 |
| 26 | R1000/pAW32 | 3 | 16.4 |
| 27 | R1000/pAW32 | 6 | 8.29 |
| 28 | R1000/pAW32 | 5 | 9.19 |
| 29 | R1000/pAW32 | 5 | 8.47 |
| 30 | R1000/pAW32 | 8 | 7.17 |
| 31 | R1000/pAW32 | 4 | 11.9 |
| 32 | R1000/pAW32 | 8 | 7.20 |
| 33 | R1000/pAW32 | 5 | 10.4 |
| 34 | R1000/pAW32 | 8 | 7.29 |
| 35 | R1000/pAW32 | 3 | 17.2 |
| 36 | R1000/pAW32 | 8 | 7.27 |
| 37 | R1000/pAW32 | 9 | 6.01 |
| 38 | R1000/pAW32 | 9 | 6.62 |
| 40 | R1000/pAW32 | 9 | 6.02 |
| 41 | R1000 | 8 | 7.23 |
| 42 | R1000 | 8 | 7.83 |
| 43 | R1000 | 10 | 6.20 |
| 44 | R1000 | 9 | 5.97 |
| 45 | R1000 | 9 | 6.73 |
| 46 | R1000 | 9 | 6.27 |
| 47 | R1000 | 8 | 7.27 |
| 48 | R1000 | 7 | 8.30 |
| 49 | R1000 | 9 | 7.11 |

**[0180]** The ability of R1000 transformed "hairy" roots to grow in the absence of exogenous phytohormones can be attributed to the Ri plasmid, pRiA4b. When R1000/pAW31 or R1000/pAW32 strains are used to transform, only a fraction (about half) of the "hairy" roots will also be transformed with the experimental binary vector, pAW31 or pAW32. Thus, as expected, not all hairy roots resulting from transformation with R1000/pAW31 show the high 18:3 phenotype.

The absense of any significant fatty acid phenotype in "hairy roots" transformed with R1000/pAW31 is expected, since carrot and Arabidopsis delta-15 desaturase sequences are not expected to be sufficiently related. These results show that overexpression of Arabidopsis microsomal delta-15 desaturase can result in over 10-fold increase in 18:3 at the expense of 18:2 in heterologous plant tissue.

Overexpression of Arabidopsis Delta-15 Fatty Acid Desaturase in Seeds and Complementation of the Mutation in Delta-15 Desaturation in Mutant 3707

**[0181]** To complement the delta-15 desaturation mutation in the T-DNA mutant 3707 and to test the biological effect of overexpression of SEQ ID NO:1 (Arabidopsis microsomal delta-15 fatty acid desaturase) in seed, the embryo-specific promoter:SEQ ID NO:1 chimeric gene was transformed into the mutant plant. This embryo-specific expression cassette in pAW42 was produced, in part, using a modified version of vector pCW109. Vector pCW109 itself was made by inserting into the Hind III site of the cloning vector pUC18 (Bethesda Research Laboratory) a 555 bp 5' non-coding region (containing the promoter region) of the β-conclycinin gene followed by the multiple cloning sequence containing the restriction endonuclease sites for Nco I, Sma I, Kpn I and Xba I, then 1174 bp of the common bean phaseolin 3' untranslated region into the Hind III site [Slightom et al., Proc. Nat'l Acad. Sci. U.S.A.(1983) 80:1897-1901]. The β-conclycinin promoter region used is an allele of the published β-conglycinin gene (Doyle et al., J. Biol. Chem. (1986) 261:9228-9238) due to differences at 27 nucleotide positions. Further sequence description may be found in Slightom (WO91/13993).

**[0182]** The modifications to vector pCW109 were as follows: The potential translation start site was destroyed by digestion with Nco I and Xba I restriction enzymes followed by treatment with mung bean nuclease (New England Biolabs) to create linear, blunt ended DNA molecules. After ligation of Not I linkers (New England Biolabs) and digestion with Not I restriction enzyme (New England Biolabs) the plasmid was recircularized. Confirmation of the desired change was obtained by dideoxy sequencing. The resulting plasmid was designated pAW35. The 1.8 kB Hind III fragment from pAW35 containing the modified β-conclycinin promotor/3' phaseolin region was subcloned into the Hind III site in plasmid vector pBluescript SK$^+$ (Stratagene) creating plasmid pAW36. Plasmid pAW36 was linerized at its unique Eco RI site and ligated to Eco RI/Xho I adaptors (Stratagene). Following digestion with Xho I, the 1.7 kB Xho I fragment containing the β-conclycinin promotor/Not I site/3'-phaseolin untranslated region was cloned into the Xho I site in pURA3 vector (Clonetech). The resultant plasmid, pAW42, contains the seed specific expression cassette bordered by Xho I sites to facilitate cloning into modified T-DNA binary vectors and a unique Not I site to facilitate cloning of target cDNA sequences. Vector pURA3 was choosen due to the absence of a Not I restriction site, the presence of a single Xho I restriction site, and the relatively large size of the vector (pURA3) would make the isolation of the gene expression cassettes relatively easy from the final construct.

**[0183]** The 1.4 kB Not I fragment in plasmid pCF3 containing Arabidopsis microsomal delta-15 desaturase (SEQ ID NO:1) was isolated and ligated to plasmid pAW42 (the seed-specific expression cassette) that had previously been linearized with Not I restriction enzyme and treated with calf intestinal alkaline phosphatase (Boehringer Mannheim) to result in plasmids pAW45 that had SEQ ID NO:1 cloned in a sense orientation with respect to the promoter. The orientation of the cDNA relative to the promotors was established by digestion with appropriate restriction endonucleases or by sequencing across the promotor-cDNA junctions.

**[0184]** The chimeric β-conclycinin promotor/sense SEQ ID NO:1/phaseolin 3' was isolated as a 3.2 kB Xho I fragment from plasmid pAW45 and subcloned into the binary vector pAW25 at its unique Sal I site. In the resulting vector, pAW50, the orientation of the plant selectable marker is the same as that of the β-conclycinin promoter as acertained by digestion with appropriate restriction endonucleases. Plasmid pAW25, is derived from plasmids p2S94K and pML2. Plasmid pZS94K contains the pBR322 origin of replication, the replication and stability regions of the Pseudomonas aeruginosa plasmid pVS1 [Itoh, et al., (1984) Plasmid 11:206-220] required for replication and maintenance of the plasmid in Agrobacterium, the bacterial NPT II gene (kanamycin resistance) from Tn5 [Berg et al., (1975) Proc. Nat'l. Acad. Sci. U.S.A. 72:3628-3632] as a selectable marker for transformed bacteria, a T-DNA left border fragment of the octopine Ti plasmid pTiA6 and right border fragment derived from TiAch5 describe by van den Elzen et al. (Plant Mol. Biol. (1985) 5:149-154). Between these borders are the E. coli lacZ a-complementing segment [Vieria and Messing (1982) Gene 19:259-267] with restriction endonuclease sites Sal I and Asp 718 derived from pUC18. A 4.5 kB Asp 718-Sal I DNA fragment containing the chimeric herbicide sulfonylurea (SU)-resistant acetolactate (ALS) gene was obtained from plasmid pML2 and cloned into the Asp 718-Sal I sites of plasmid pZS94K. This chimeric ALS gene contained the CaMV 35S promoter/Cab22L Bgl II-Nco I fragment that is described by Harpster et al., [Mol. Gen. Genet. (1988) 212: 182-190] and the Arabidopsis ALS coding and 3' non-coding sequences (Mazur et al., (1987) Plant Physiol. 85; 1110-1117] that was mutated so that it encodes a SU-resistant form of ALS. The mutation, introduced by site-directed mutagenesis, are those present in the tobacco SU-resistant Hra gene described by Lee et al., (1988) EMBO J. 5: 1241-1248. The resulting plasmid was designated pAW25.

**[0185]** The binary vector pAW25 containing the chimeric embryo-specific β-conglycinin promotor:sense SEQ ID NO:

1 gene was transformed by the freeze/thaw method [Holsters et al., (1978) Mol. Gen. Genet. 163:181-187] into the avirulent Agrobacterium strain LBA4404/pAL4404 [Hoekema et al., (1983) Nature 303:179-180).

**[0186]** Arabidopsis root cultures were transformed by co-cultivation with Agrobacterium using standard aseptic techniques for the manipulation of sterile media and axenic plant/bacterial cultures were followed, including the use of a laminar flow hood for all transfers. Compositions of the culture media are listed in Table 8. Unless otherwise indicated, 25x100 mm petri plates were used for plant tissue cultures. Incubation of plant tissue cultures was at 23°C under constant illumination with mixed fluorescent and "Gro and Sho" plant lights (General Electric) unless otherwise noted. To initiate in vitro root cultures of the T-DNA homogyzous mutant line 3707 (Arabidopsis thaliana (L.) Heynh, geographic race Wassilewshija) seeds of the mutant line were sterilized for 10 min in a solution of 50% Chlorox with 0.1% SDS, rinsed 3 to 5 times with sterile $dH_2O$, dried thoroughly on sterile filter paper, and then 2-3 seeds were sown in liquid B5 medium in 250 mL Belco flasks. The flasks were capped, placed on a rotary shaker at 70-80 rpm, and incubated for 3-4 weeks. Prior to inoculation with Agrobacterium, root tissues were cultured on callus induction medium (MSKig). Roots were harvested by removing the root mass from the Belco flask, placing it in a petri dish, and, using forceps, pulling small bundles of roots from the root mass and placing them on MSKig medium. Petri dishes were sealed with filter tape and incubated for four days.

**[0187]** Agrobacterium strain LBA4404 carrying the plasmids pAL4404 and pAW50 were grown in 5 mL of YEB broth containing 25 mg/L kanamycin and 100 mg/L rifampicin. The culture was grown for approximately 17-20 h in glass culture tubes in a New Brunswick platform shaker (225 rpm) maintained at 28°C. Pre-cultured roots were cut into 0.5 cm segments and placed in a 100 µm filter, made from a Tri-Pour beaker (VWR Scientific, San Francisco, CA USA) and wire mesh, which is set in a petri dish. Root segments were inoculated for several min in 30-50 mL of a 1:20 dilution of the overnight Agrobacterium culture with periodic gentle mixing. Inoculated roots were transferred to sterile filter paper to draw off most of the liquid. Small bundles of roots, consisting of several root segments, were placed on MSKig medium containing 100 µM acetosyringone (3',5'-Dimethoxy-4'-hydroxyaceto-phenone, Aldrich Chemical Co., Milwaukee, WI, USA). Petri plates were sealed with parafilm or filter tape and incubated for 2 to 3 days.

**[0188]** After infection, root segments were rinsed and transferred to shoot induction medium with antibiotics. Root bundles were placed in a 100 µm filter unit (described above) and rinsed with 30-50 mL liquid MSKig medium. The filter was vigorously shaken in the solution to help remove the Agrobacterium, transferred to a clean petri dish, and rinsed again. Roots were blotted on sterile filter paper and bundles of roots were placed on MSg medium containing 500 mg/l vancomycin and either 10 or 20 ppb chlorsulfuron. Plates were sealed with filter tape and incubated for 12 to 14 days.

**[0189]** Green nodules and small shoot primordia were visible at about 2-3 weeks. The explants were either left intact or were broken into numerous pieces and placed on GM medium containing 200-300 mg/L vancomycin and either 10 or 20 ppb chlorsulfuron for further shoot development. Plates were either sealed with two pieces of tape or with filter tape. As they developed, individual shoots were isolated from the callus and were placed on MSRg medium containing 100 mg/L vancomycin and either 10 or 20 ppb chlorsulfuron. Dishes were sealed as described above and incubated for seven to 10 days. Shoots were then transferred to GM medium containing 100-200 mg/L vancomycin in 25x100 petri dishes or Magenta G7 vessels. Many primary transformants (T1) which were transferred to individual containers set seed (T2).

**[0190]** T2 seed was harvested from selected putative transformants and sown on GM medium containing 10ppb chlorsulfuron. Plates were sealed with filter tape, cold treated for 2 or more days at 4°C, and then incubated for 10 to 20 days at 23°C under constant illumination as described above. Seedlings were scored as resistant (green, true leaves develop) and sensitive (no true leaves develop).

**[0191]** Selected chlorsulfuron resistant T2 seedlings were transplanted to soil and were grown to maturity at 23°C daytime (16 h) 18°C nighttime (8 h) at 65-80% relative humidity.

**[0192]** T2 seeds from two plants were harvested at maturity and analysed individually for fatty acid composition by gas chromatography of the fatty acyl methyl esters essentially as described by Browse et al., (Anal. Biochem. (1986) 152:141-145) except that 2.5% $H_2SO_4$ in methanol was used as the methylation reagent and samples were heated for 1.5 h at 80°C to effect the methanolysis of the seed triglycerides. The results are shown in Table 7.

TABLE 7

| Percent Fatty Acid in Seeds of Transgenic Mutant 3707 | | | | | |
|---|---|---|---|---|---|
| Seed Sample | 16:0 | 18:0 | 18:1 | 18:2 | 18:3 |
| wildtype(6) | 6 | 4 | 14 | 30 | 19 |
| mutant 3707(6) | 6 | 4 | 14 | 44 | 3 |
| 1-1 | 10 | 4 | 22 | 9 | 55 |
| 1-2 | 11 | 6 | 22 | 14 | 48 |

TABLE 7 (continued)

| Percent Fatty Acid in Seeds of Transgenic Mutant 3707 | | | | | |
|---|---|---|---|---|---|
| Seed Sample | 16:0 | 18:0 | 18:1 | 18:2 | 18:3 |
| 1-3 | 12 | 7 | 16 | 6 | 57 |
| 1-4 | 10 | 4 | 30 | 52 | 4 |
| 1-5 | 10 | 4 | 18 | 17 | 48 |
| 1-6 | 10 | 5 | 15 | 15 | 53 |
| 2-1 | 11 | 5 | 19 | 60 | 4 |
| 2-2 | 10 | 5 | 19 | 9 | 56 |
| 2-3 | 9 | 4 | 27 | 8 | 52 |
| 2-4 | 10 | 5 | 17 | 10 | 56 |
| 2-5 | 10 | 5 | 19 | 9 | 56 |
| 2-6 | 10 | 5 | 17 | 17 | 48 |

**[0193]** The fatty acid composition of the wild-type and mutant line 3707 represents the average of 6 single seeds each. Seeds from plant 1 are designated 1-1 to 1-6 and those from plant 2 are designated 2-1 to 2-6. The 20:1 and 20:2 amounts are not shown. The data shows that the one out of six seeds in each plant show the mutant fatty acid phenotype, while the remaining seeds show more than 10-fold increase in 18:3 to ca.55%. While most of the increase occurs at the expense of 18:2, some of it also occurs at the expense of 18:1. Such high levels are of linolenic acid in vegetable oils are observed in specialty oil crops, such as linseed. Thus, overexpression of this gene in other oilscrops, especially canola, which is a close relative of Arabidopsis, is also expected to result in such high levels of 18:3.

TABLE 8

| Medium Composition | | |
|---|---|---|
| YEP MEDIUM | | BASIC MEDIUM |
| Bacto Beef Extract | 5.0 g | 1 Pkg. Murashige and Skoog Minimal Organics Medium without Sucrose (Gibco #510-3118 or Sigma #M6899) |
| Bacto Yeast Extract | 1.0 g | |
| Peptone | 5.0 g | |
| Sucrose | 5.0 g | 10 mL Vitamin Supplement |
| $MgSO_4 \cdot 7H_2O$ | 0.5 g | 0.05% MES 0.5 g/L |
| Agar (optional) | 15.0 g | 0.8% agar 8 g/L |
| pH | | pH |
| VITAMIN SUPPLEMENT | | GM = Germination Medium |
| 10 mg/L thiamine | | Basic Medium |
| 50 mg/L pyridoxine | | 1% sucrose 10 g/L |
| 50 mg/L nicotinic acid | | |
| MSKlg = Callus Induction Medium MSg = Shoot Induction Medium | | |
| Basic Medium | | Basic Medium |
| 2% glucose | 20 g/L | 2% glucose 20 g/L |
| 0.5 mg/L 2,4-D | 2.3 µL | 0.15 mg/L IAA 0.86 µM |
| 0.3 mg/L Kinetin | 1.4 µM | 5.0 mg/L 2iP 24.6 µM |
| 5 mg/L IAA | 28.5 µM | |
| MSRg = Shoot Induction Medium | | |
| Basic Medium | | |
| 2% glucose | 20 g/L | |
| 12 mg/L IBA | 58.8 µM | |
| 0.1 mg/L Kinetin | 0.46 µM | |

EXAMPLE 10

Construction of Vectors for Transformation of Brassica napus for Reduced Expression of Delta-15 Desaturases in Developing Seeds

**[0194]** Detailed procedures for manipulation of DNA fragments by restriction endonuclease digestion, size separation by agarose gel electrophoresis, isolation of DNA fragments from agarose gels, ligation of DNA fragments, modification of cut ends of DNA fragments and transformation of E. coli cells with circular DNA plasmids are all described in Sambrook et al., (Molecular Cloning, A Laboratory Manual, 2nd ed (1989) Cold Spring Harbor Laboratory Press) and Ausubel et al., Current Protocols in Molecular Biology (1989) John Wiley & Sons).

**[0195]** Sequences of the cDNA's encoding the B. napus cytoplasmic delta-15 desaturase and the Brassica napus plastid delta-15 desaturase were placed in the antisense orientation behind the promoter region from the a' subunit of the soybean storage protein β-conglycinin to provide embryo specific expression and high expression levels.

**[0196]** An embryo-specific expression cassette was constructed to serve as the basis for chimeric gene constructs for anti-sense expression of the nucleotide sequences of delta-15 desaturase cDNAs. The vector pCW109 was produced by the insertion of 555 base pairs of the β-conglycinin (a' subunit of the 7s seed storage protein) promoter from soybean (Glycine max), the β-conglycinin 5' untranslated region followed by a multiple cloning sequence containing the restriction endonuclease sites for Nco I, Sma I, Kpn I and Xba I, then 1174 base pairs of the common bean phaseolin 3' untranslated region into the Hind III site in the cloning vector pUC18 (BRL). The β-conglycinin promoter sequence represents an allele of the published β-conglycinin gene (Doyle et al., (1986) J. Biol. Chem. 261:9228-9238) due to differences at 27 nucleotide positions. Further sequence description may be found in Slightom (WO91/13993) The sequence of the 3' untranslated region of phaseolin is described in (Slightom et al., (1983) Proc. Natl. Acad. Sci. USA, 80:1897-1901).

**[0197]** To facilitate use in antisense constructions, the Nco I site and potential translation start site in the plasmid pCW109 was destroyed by digestion with Nco I, mung bean exonuclease digestion and re-ligation of the blunt site to give the modified plasmid pCW109A. pCW109A was opened between the β-conglycinin promoter sequence and the phaseolin 3' sequence by digestion with Sma I to allow insertion of blunt ended cDNA fragments encoding the delta-15 desaturase sequences by ligation. The blunt ended fragment of the cytoplasmic delta-15 desaturase was obtained from plasmid pBNSF3, which contains the nucleotides 208 to 1336 of the cDNA insert described in SEQ ID NO:6. pBNSF3 was modified to remove the Hind III site at bases 682 to 687 of SEQ ID 6 by digesting with Hind III, blunting with Klenow and religating. The resulting plasmid [pBNSF3(-H)], was digested with Eco RI and Xho I to release the delta-15 cDNA fragment, all ends were Klenow blunted and the 1.2 kB coding region was purified by gel isolation. The 1.2 kB fragment was ligated into the Sma I cut pCW109A described above. The antisense orientation of the inserted cDNA relative to the β-conglycinin promoter was established by digestion with Aat I which cuts in the delta-15 desaturase coding region and in the vector 5' to the β-conglycinin promoter to release a 1.4 Kb fragment when the coding region is in the antisense orientation. The antisense construction was given the name pCCFdR1.

**[0198]** The transcription unit [β-conglycinin promoter:antisense delta-15 desaturase:phaseolin 3'end] was released from pCCFdR1 by Hind III digestion, isolated,and ligated into pBluescript which had also been Hind III digested to give plasmid pCCFdR2. This construct has unique BamH I and Sal I sites which were digested. The 3 kB transcriptional unit was isolated and cloned into the Bam HI and Sal I sites in p2199 described below to give the binary vector pZCC3FdR. The orientation given by this directional cloning is with transcription of both the selectable marker gene and the delta-15 antisense gene in the same direction and toward the right border tDNA sequence.

**[0199]** An antisense construction based on the plastid delta-15 desaturase was made with the 425 most 3' bases of SEQ ID NO:8 which is contained in the plasmid pBNSFD-8. pBNSFD-8 represents a cDNA of the plastid delta-15 desaturase in pBluescript. The cDNA insert was removed from pBNSFD-8 by digestion with Xho I and Sma I, the fragments were blunted, and the 425 base insert isolated by gel purification. The isolated fragment was cloned into the Sma I site of pCW109A and the antisense orientation of the chosen clone confirmed by digestion of the plasmid with Pst I. Pst I cuts in the plastid delta-15 sequence and in the pCW109A vector 5' to the β-conglycinin promoter to release a 1.2 kB fragment indicative of the antisense orientation. The plasmid containing this construction was called pCCdFdR1.

**[0200]** Digestion of pCCdFdR1 with Hind III removes a 2.3 kB fragment containing the transcriptional unit [β-conglycinin promter:plastid delta-15 antisense:3'-phaseolin sequence]. The fragment was gel isolated and cloned into Hind III digested pBluescript. The orientation of the fragment was relative to the Bam HI site in the cloning region of pBluescript was determined by digestion with Pst I as described above. A clone oriented with the promoter toward the Sal I containing end was chosen and given the name pCCdFdR2.

**[0201]** pCCdFdR2 was digested with Bam HI and Sal I, the released fragment was gel isolated and ligated into pZ199 which had been digested with Bam HI and Sal I to give the binary vector pZCCdFdR.

**[0202]** Vectors for transformation of the antisense delta-15 desaturase constructions under control of the β-congly-

cinin promoter into plants using Agrobacterium tumefaciens were produced by constructing a binary Ti plasmid vector system (Bevan, (1984) Nucl. Acids Res. 12:8711-8720). The starting vector used for these systems (pZS199) is based on a vector which contains: (1) the chimeric gene nopaline synthase/neomycin phosphotransferase as a selectable marker for transformed plant cells (Bevan et al., (1984) Nature 304:184-186), (2) the left and right borders of the T-DNA of the Ti plasmid (Bevan et al., (1984) Nucl. Acids Res. 12:8711-8720), (3) the E. coli lacZ a-complementing segment (Vieria and Messing (1982) Gene 19:259-267) with unique restriction endonuclease sites for Eco RI, Kpn I, Bam HI, Hin DIII, and Sal I, (4) the bacterial replication origin from the Pseudomonas plasmid pVS1 (Itoh et al., (1984) Plasmid 11:206-220), and (5) the bacterial neomycin phosphotransferase gene from Tn5 (Berg et al., (1975) Proc. Natnl. Acad. Sci. U.S.A. 72:3628-3632) as a selectable marker for transformed A. tumefaciens. The nopaline synthase promoter in the plant selectable marker was replaced by the 35S promoter (Odell et al. (1985) Nature, 313:810-813) by a standard restriction endonuclease digestion and ligation strategy. The 35S promoter is required for efficient Brassica napus transformation as described below.

## EXAMPLE 11

## AGROBACTERIUM MEDIATED TRANSFORMATION OF BRASSICA NAPUS

**[0203]** The binary vectors pZCC3FdR abd pZCCdFdR were transferred by a freeze/thaw method (Holsters et al., (1978) Mol Gen Genet 163:181-187) to the Agrobacterium strain LBA4404/pAL4404 (Hoekema et al., (1983), Nature 303:179-180).

**[0204]** Brassica napus cultivar "Westar" was transformed by co-cultivation of seedling pieces with disarmed Agrobacterium tumefaciens strain LBA4404 carrying the the appropriate binary vector.

**[0205]** B. napus seeds were sterilized by stirring in 10% Chlorox, 0.1% SDS for thirty min, and then rinsed thoroughly with sterile distilled water. The seeds were germinated on sterile medium containing 30 mM $CaCl_2$ and 1.5% agar, and grown for six days in the dark at 24°C.

**[0206]** Liquid cultures of Agrobacterium for plant transformation were grown overnight at 28°C in Minimal A medium containing 100 mg/L kanamycin. The bacterial cells are pelleted by centrifugation and resuspended at a concentration of $10^8$ cells/mL in liquid Murashige and Skoog Minimal Organic medium containing 100 µM acetosyringone.

**[0207]** B. napus seedling hypocotyls were cut into 5 mm segments which were immediately placed into the bacterial suspension. After 30 min, the hypocotyl pieces were removed from the bacterial suspension and placed onto BC-12 callus medium containing 100 µM acetosyringone. The plant tissue and Agrobacteria were co-cultivated for three days at 24°C in dim light.

**[0208]** The co-cultivation was terminated by transferring the hypocotyl pieces to BC-12 callus medium containing 200 mg/L carbenicillin to kill the Agrobacteria, and 25 mg/L kanamycin to select for transformed plant cell growth. The seedling pieces were incubated on this medium for three weeks at 24°C under continuous light.

**[0209]** After three weeks, the segments wre transferred to BS-48 regeneration medium containing 200 mg/L carbenicillin and 25 mg/L kanamycin. Plant tissue was subcultured every two weeks onto fresh selective regeneration medium, under the same culture conditions described for the callus medium. Putatively transformed calli grow rapidly on regeneration medium; as calli reached a diameter of about 2 mm, they were removed from the hypocotyl pieces and placed on the same medium lacking kanamycin.

**[0210]** Shoots began to appear within several weeks after transfer to BS-48 regeneration medium. As soon as the shoots formed discernable stems, they were excised from the calli, transferred to MSV-1A elongation medium, and moved to a 16:8 h day/night photoperiod at 24°C.

**[0211]** Once shoots had elongated several internodes, they were cut above the agar surface and the cut ends were dipped in Rootone. Treated shoots were planted directly into wet Metro-Mix 350 soiless potting medium. The pots were covered with plastic bags which were removed when the plants were clearly growing -- after about 10 days.

**[0212]** Plants were grown under a 16:8 h day/night photoperiod, with a daytime temperature of 23°C and a nighttime temperature of 17°C. When the primary flowering stem began to elongate, it was covered with a mesh pollen-containment bag to prevent outcrossing. Self-pollination was facilitated by shaking the plants several times each day. Seeds derived from self-pollinations were harvested about three months after planting.

TABLE 9

| Minimal A Bacterial Growth Medium | Brassica callus Medium BC-12 |
|---|---|
| Dissolve in distilled water: | Per liter: |
| 10.5 g potassium phosphate, dibasic | Murashige and Skoog Minimal Organic Medium (MS salts, 100 |
| 4.5 g potassium phosphate, monobasic | mg/L i-inositol, 0.4 mg/L |
| 1.0 g ammonium sulfate | thiamine; GIBCO #510-3118) |

TABLE 9 (continued)

| | |
|---|---|
| 0.5 g sodium citrate, dihydrate<br>Make up to 979 mLs with<br>distilled water<br>Autoclave<br>Add 20 mLs filter-sterilized 10% sucrose<br>Add 1 mL filter-sterilized<br>1 M $MgSO_4$ | 30 sucrose<br>18 g mannitol<br>1.0 mg/L 2,4-D<br>3.0 mg/L kinetin<br>0.6% agarose<br>pH 5.8 |
| Brassica Regeneration Medium BS-48<br>Murashige and Skoog Minimal<br>Organic Medium Gamborg B5 Vitamins<br>(SIGMA #1019)<br>10 g glucose<br>250 mg xylose<br>600 mg MES<br>0.4% agarose<br>pH 5.7<br>Filter-sterilize and add after<br>autoclaving:<br>2.0 mg/L zeatin<br>0.1 mg/L IAA | Brassica Shoot Elongation<br>Medium MSV-1A<br>Murashige and Skoog Minimal<br>Organic Medium Gamborg B5<br>vitamins<br>10 g sucrose<br>0.6% agarose<br>pH 5.8 |

EXAMPLE 12

ANALYSIS OF TRANSGENIC BRASSICA NAPUS PLANTS

**[0213]** Insertion of the intact antisense transcriptional unit was verified by Southern analysis using transgenic plant leaf tissue as the source of DNA as described in Example 5. Ten micrograms of leaf DNA was digested to completion with a mixture of Bam HI and Sal I restriction endonucleases and then separatd by agarose gel electrophoresis. The separated DNA was transferred to Hybond $H^+$ membrane and hybridized with radiolabeled insert from pBNSF3-2. An estimate of the number of copies of the inserted transgene was made by calibrating each Southern blot with standard amounts of pBNSF3-2 corresponding to 1 and 5 copies per genome and comparing intensities of the autoradiographic signal from the standards, the endogenous delta-15 desaturase signals and the inserted gene signal. To date, 38 independent transformants have been analyzed for presence of the gene and 36 were found to be positive.

**[0214]** The relative content of the 5 most abundant fatty acids in canola seeds was determined either by direct trans-esterification of individual seeds in 0.5 mL of methanolic $H_2SO_4$ (2.5%) or by hexane extraction of bulk seed samples followed by trans-esterification of an aliquot in 0.8 mL of 1% sodium methoxide in methanol. Fatty acid methyl esters were extracted from the methanolic solutions into hexane after the addition of an equal volume of water.

**[0215]** The relative content of 18:3 fatty acid varies significantly during seed development. To a lesser extent, the ratio of 18:3 to 18:2 varies also. Thus meaningful data can be obtained only from seeds after maturation and drydown. Additionally, the ratio of 18:3 to total fatty acid content and to 18:0 varies significantly due to environmental factors, primarily temperature. In this circumstance, the most appropriate controls are the transformed plants which by Southern analysis do not contain the antisense delta-15 transgene. Analysis from the first 5 transformants to reach dry seed are given in Table 10 below. Seeds were harvested using a hand thesher, bulked and a 1.5 g (about 300 seeds) sample was taken. Seed from each transformant was crushed with a mortar and pestel, extracted 4 times with 8 mL hexane at about 50°C. The combined extracts were reduced in volume to 5 mL and two 50 microliter aliquots were taken for esterification as described above. Separation of the fatty acid methyl esters was done by gas-liquid chromatography using an Omegawax 320 column (Supelco Inc., 0.32 mm ID X 30M) run isothermally at 220° and cycled to 260° between each injection.

TABLR 10

| Transformant No. | % 18:3 | %18:3/18:2 | Antisense delta-15 Copy No. |
|---|---|---|---|
| pZCC3FdR-91 | 6.2 | 0.39 | 0 |
| pZCC3FdR-81 | 5.9 | 0.33 | 1 |

TABLR 10 (continued)

| Transformant No. | % 18:3 | %18:3/18:2 | Antisense delta-15 Copy No. |
|---|---|---|---|
| pZCC3FdR-15 | 6.0 | 0.38 | 2 |
| pZCC3FdR-11 | 5.6 | 0.34 | 1 |
| pZCC3FdR-148 | 8.2 | 0.40 | 2 |

**[0216]** The differences between the 4 transformed lines and line 92 are very small, however to test the significance of the difference in the 18:3/18:2 ratio between line 81 and 91, 25 individual seeds from each line were transesterified and their fatty acid composition determined. The average ratio for line 81 was 0.345 with a coefficient of variation of 11.6% while the average for line 91 was 0.375 with a coefficient of variation of 8.0%. The sample means are significantly different at the 0.01% level using Student's t test.

EXAMPLE 13

CONSTRUCTION OF VECTORS FOR TRANSFORMATION OF GLYCTNE MAX FOR REDUCED EXPRESSION OF DELTA-15 DESATURASES IN DEVELOPING SEEDS

**[0217]** The antisense G. max plastid delta-15 desaturase CDNA sequence under control of the β-conglycinin promoter was constructed using the vector pCW109A described in Example 10 above. For use in the soybean transformation system described below, the transcriptional unit was placed in a vector along with an appropriate selectable marker expression system. The starting vector was pML45, which consists of the non-tissue specific and constitutive promoter designated 508D and described in Hershey (WO 9011361) driving expression of the neomycin phosphotransferase gene described in (Beck et al. (1982) Gene 19:327-336) followed by the 3' end of the nopaline synthase gene including nucleotides 848 to 1550 described by (Depicker et al. (1982) J. Appl. Genet. 1:561-574). This transcriptional unit was inserted into the commercial cloning vector pGEM9Z (BRL) and is flanked at the 5' end of the 508D promoter by the restriction sites Sal I, Xba I, Bam HI and Sma I in that order. An additional Sal I site is present at the 3' end of the NOS 3' sequence and the Xba I, Bam HI and Sal I sites are unique.

**[0218]** Removal of the unit [β-conglycinin promter:cloning region:phaseolin 3' end] from pCW109A by digestion with Hind III, blunting the ends and isolating the 1.8 kB fragment afforded the expression cassette pCST by ligating the above isolated fragment into the Sma I site of pML45. A clone with the β-conglycinin promoter in the same orientation as the 508D promoter were chosen by digestion with Xba I. The correct orientation releases a 700 bp fragment. This vector cassette was called pCST.

**[0219]** The 2.2 kB insert encoding the soybean, plastid delta-15 desaturase was subcloned from the plasmid pXF1 by digestion with HinP I to remove about 1 kB of unrelated cDNA. HinP I cuts within the cDNA insert very near the 5' end of the cDNA for the delta-15 desaturase and about 300 bp from the 3' end of that cDNA. The Cla I compatable ends were cloned into Cla I digested pBluescript and a clone with the 5' end of the cDNA toward the Eco RV site in the pBluescript cloning region was selected based on the relaese of a 900 bp fragment by digestion with Pst I. The subcloned plasmid was called pS3Fd1.

**[0220]** The delta-15 encoding sequence was removed from pS3Fd1 by digestion with HinC II and Eco RV, the 2.2 kB fragment was gel isolated and cloned into the opened Sma I site in pCST1. A clone with the delta-15 sequence in the antisense orientation to the β-conglycinin promoter was selected by digestion with Xba I. The antisense construct releases a 400 bp piece and that clone was designated pCS3FdST1R.

EXAMPLE 14

TRANSFORMATION OF SOMATIC SOYBEAN EMBRYO CULTURES

**[0221]** Soybean embryogenic suspension cultures are maintained in 35 mL liquid media (SB55 or SBP6) on a rotary shaker, 150 rpm, at 28°C with mixed florescent and incandescent lights on a 16:8 h day/night schedule. Cultures were subcultured every four weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

**[0222]** Soybean embryogenic suspension cultures were transformed with pCS3FdST1R by the method of particle gun bombardment (see Kline et al. (1987) Nature (London) 327:70). A Du Pont Biolistic PDS1000/HE instrument (helium retrofit) was used for these transformations.

**[0223]** To 50 mL of a 60 mg/mL 1 mm gold particle suspension was added (in order); 5 µL DNA(1 µg/µL), 20 µL spermidine (0.1M), and 50 µl $CaCl_2$ (2.5 M). The particle preparation was agitated for 3 min, spun in a microfuge for 10 sec and the supernatant removed. The DNA-coated particles were then washed once in 400 µL 70% ethanol and

resuspended in 40 µL of anhydrous ethanol. The DNA/particle suspension was sonicated three times for 1 sec each. Five µL of the DNA-coated gold particles were then loaded on each macro carrier disk.

**[0224]** Approximately 300-400 mg of a four week old suspension culture was placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue were normally bombarded. Membrane rupture pressure was set at 1000 psi and the chamber was evacuated to a vacuum of 28 inches of mercury. The tissue was placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue was placed back into liquid and cultured as described above.

**[0225]** Eleven days post bombardment, the liquid media was exchanged with fresh SB55 containing 50 mg/mL hygromycin. The selective media was refreshed weekly. Seven weeks post bombardment, green, transformed tissue was observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue was removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Thus each new line was treated as independent transformation event. These suspensions can then be maintained as suspensions of embryos clustered in an immature developmental stage through subculture or regenerated into whole plants by maturation and germination of individual somatic embryos.

**[0226]** Transformed embryogenic clusters were removed from liquid culture and placed on a solid agar media (SB103) containing no hormones or antibiotics. Embryos were cultured for eight weeks at 26°C with mixed florescent and incandescent lights on a 16:8 h day/night schedule. During this period, individual embryos were removed from the clusters and analyzed at various stages of embryo development After eight weeks the embryos become suitable for germination.

TABLE 11

| Media: | | B5 Vitimin Stock |
|---|---|---|
| SB55 and SBP6 Stock Solutions | | 10 g m-inositol |
| (g/L): | | 100 mg nicotinic acid |
| MS Sulfate 100X Stock | | 100 mg pyridoxine HCl |
| $MgSO_4$ 7H2O | 37.0 | 1 g thiamine |
| $MnSO_4$ H2O | 1.69 | SB55 (per Liter) |
| $ZnSO_4$ 7H2O | 0.86 | 10 mL each MS stocks |
| $CuSO_4$ 5H2O | 0.0025 | 1 mL B5 Vitaimin stock |
| MS Halides 100X Stock | | 0.8 g $NH_4NO_3$ |
| $CaCl_2$ $2H_2O$ | 44.0 | 3.033 g $KNO_3$ |
| KI XI | 0.083 0083 | 1 mL 2,4-D (10mg/mL stock) |
| $CoCl_2$ $6H_2O$ | 0.00125 | 60 g sucrose |
| $KH_2PO_4$ | 17.0 | 0.667 g asparagine |
| $H_3BO_3$ | 0.62 | pH 5.7 |
| $Na_2MoO_4$ $2H_2O$ | 0.025 | For SBP6- substitute 0.5 mL 2,4-D |
| MS FeEDTA 100X Stock | | SB103 (per Liter) |
| $Na_2EDTA$ | 3.724 | MS Salts |
| $FeSO_4$ $7H_2O$ | 2.784 | 6% maltose |
| | | 750 mg $MgCl_2$ |
| | | 0.2% Gelrite |
| | | pH 5.7 |

## EXAMPLE 15

## ANALYSIS OF TRANSGENIC GLYCINE MAX PLANTS

**[0227]** While in the globular embryo state in liquid culture as described in Example 14, somatic soybean embryos contain very low amounts of triacylglycerol or storage proteins typical of maturing, zygotic soybean embryos. At this developmental stage, the ratio of total triacylglyceride to total polar lipid (phospholipids and glycolipid) is about 1:4, as is typical of zygotic soybean embryos at the developmental stage from which the somatic embryo culture was initiated. At the globular stage as well, the mRNAs for the prominant seed proteins (a' subunit of β-conglycinin, Kunitz Trypsin Inhibitor III and Soybean Seed Lectin) are essentially absent. Upon transfer to hormone free media to allow differen-

tiation to the maturing somatic embryo state as described in Example 14, triacylglycerol becomes the most abundant lipid class. As well, mRNAs for a'-subunit of β-conglycinin, Kunitz Trypsin Inhibitor III and Soybean Seed Lectin become very abundant messages in the total mRNA population. In these respects the somatic soybean embryo system behaves very similarly to maturing zygotic soybean embryos in vivo, and is therefore a good and rapid model system for analyzing the phenotypic effects of modifying the expression of genes in the fatty acid biosynthesis pathway. Similar somatic embryo culture systems have been documented and used in another oilseed crop, rapeseed (Taylor et al. (1990) Planta 181:18-26). Fatty acid analysis was performed as described in Example 12 using single embryos as the tissue source. A number of embryos from line 2872 (control tissue transformed with pCST) and lines 299,303,306 and 307 (line 2872 transformed with plasmid pCS3FdST1R) were analyzed for fatty acid content. The relative fatty-acid composition of embryos taken from tissue transformed with pCS3FdST1R was compared with control tissue, transformed with pCST. The results of this analysis are shown in Table 12.

TABLE 12

| Line | Embryo | 16:0 | 18:0 | 18:1 | 18:2 | 18:3 |
|---|---|---|---|---|---|---|
| 2872 | 1 | 17.7 | 4.1 | 11.3 | 52.8 | 14.1 |
| | 2 | 17.3 | 4.3 | 10.9 | 49.5 | 18.0 |
| | 3 | 16.1 | 4.1 | 13.8 | 48.2 | 17.3 |
| | 4 | 17.5 | 3.6 | 11.7 | 52.0 | 14.1 |
| | 5 | 16.6 | 3.9 | 12.7 | 53.7 | 12.6 |
| | 6 | 14.8 | 3.0 | 14.7 | 55.3 | 11.1 |
| | av | 16.7 | 3.8 | 12.5 | 51.9 | 14.5 |
| 299-1-3 | 1 | 16.5 | 4.1 | 9.7 | 61.4 | 6.3 |
| 299-15-1 | 1 | 14.7 | 3.6 | 11.9 | 61.3 | 8.4 |
| | 2 | 16.6 | 3.7 | 12.1 | 58.6 | 8.6 |
| | 3 | 16.7 | 4.1 | 14.9 | 53.2 | 11.1 |
| | 4 | 15.2 | 4.0 | 9.1 | 60.2 | 11.5 |
| | 5 | 16.0 | 4.2 | 13.9 | 55.2 | 10.7 |
| | 6 | 15.2 | 3.5 | 9.9 | 63.4 | 8.1 |
| 303-7-1 | 1 | 14.1 | 2.2 | 10.6 | 59.4 | 13.7 |
| | 2 | 14.0 | 2.8 | 12.5 | 59.3 | 11.4 |
| 306-4-5 | 1 | 17.5 | 4.2 | 8.1 | 62.7 | 7.4 |
| | 215.7 | 3.3 | 9.0 | 60.5 | 11.5 | |
| | 3 | 17.1 | 3.4 | 9.3 | 60.7 | 9.5 |
| | 4 | 15.7 | 3.8 | 9.2 | 61.2 | 9.7 |
| | 5 | 17.7 | 3.9 | 6.5 | 58.3 | 13.6 |
| | 6 | 16.6 | 3.4 | 10.2 | 59.2 | 10.6 |
| 306-4-8 | 1 | 16.6 | 3.9 | 15.3 | 50.7 | 11.8 |
| | 2 | 17.8 | 3.6 | 15.7 | 50.0 | 10.8 |
| | 3 | 16.7 | 3.3 | 11.1 | 52.0 | 14.6 |
| | 4 | 19.0 | 4.0 | 10.3 | 53.1 | 12.3 |
| | 5 | 19.7 | 3.5 | 9.0 | 53.6 | 13.0 |
| | 6 | 18.0 | 2.9 | 13.1 | 52.8 | 10.9 |
| 307-1-1 | 1 | 14.4 | 3.7 | 11.2 | 64.4 | 6.3 |
| | 2 | 15.4 | 3.4 | 7.8 | 61.0 | 11.3 |
| | 3 | 17.2 | 2.5 | 12.0 | 57.2 | 11.1 |
| 307-1-2 | 1 | 13.4 | 3.0 | 8.4 | 55.4 | 19.9 |
| | 2 | 16.3 | 3.1 | 6.4 | 55.7 | 18.7 |
| | 3 | 14.0 | 3.3 | 8.8 | 58.7 | 15.2 |
| | 4 | 15.8 | 2.5 | 9.8 | 59.7 | 12.2 |
| | 5 | 14.6 | 3.7 | 14.9 | 51.1 | 15.7 |
| | 6 | 14.3 | 3.9 | 11.4 | 55.5 | 14.1 |
| 307-1-3 | 1 | 14.8 | 3.1 | 9.4 | 60.5 | 12.2 |
| | 2 | 18.0 | 3.0 | 5.3 | 56.2 | 15.2 |

TABLE 12 (continued)

| Line | Embryo | 16:0 | 18:0 | 18:1 | 18:2 | 18:3 |
|---|---|---|---|---|---|---|
| | 3 | 18.0 | 3.4 | 2.5 | 58.6 | 15.4 |
| 307-1-4 | 1 | 15.0 | 2.7 | 13.8 | 61.7 | 6.9 |
| | 2 | 15.9 | 2.7 | 9.8 | 62.0 | 9.6 |
| | 3 | 14.6 | 3.2 | 13.4 | 61.4 | 6.7 |
| 307-1-5 | 1 | 15.9 | 3.5 | 7.6 | 61.7 | 11.2 |
| | 2 | 14.6 | 3.5 | 10.0 | 61.3 | 10.6 |
| | 3 | 18.7 | 2.6 | 6.8 | 53.0 | 19.0 |
| 307-1-7 | 1 | 15.3 | 3.5 | 12.5 | 60.3 | 8.5 |
| | 2 | 16.2 | 2.2 | 13.9 | 57.1 | 10.6 |
| | 3 | 14.9 | 3.1 | 12.2 | 58.0 | 11.8 |
| 307-1-9 | 1 | 16.4 | 2.9 | 23.2 | 47.9 | 9.6 |
| | 2 | 19.6 | 0.0 | 20.4 | 51.3 | 8.8 |
| | 3 | 16.8 | 3.3 | 24.6 | 49.6 | 5.7 |
| 307-1-11 | 1 | 18.1 | 3.6 | 5.7 | 52.9 | 19.7 |
| | 2 | 14.7 | 3.7 | 9.9 | 58.7 | 13.0 |
| | 3 | 15.1 | 3.7 | 11.3 | 55.8 | 14.1 |

**[0228]** The average 18:3 content of control embryos was 14.5% with a range from 11.1% to 18.0%. The average 18:3 content of transformed embryos was 11.5% with a range of 6.3% to 19.9%. Almost 80% of the transformed embryos (38/48) had an 18:3 content below that of the control mean. About 44% had an 18:3 content less than the lowest observed control value and 12.5% had an 18:3 content less than half of the control mean value (i.e., less than 7.5%). The lowest 18:3 content observed in transformed tissue was 6.3% (299-1-3, 307-1-2 #1) compared with the control low of 11.1%. In all cases in transformed tissue, a decrease in 18:3 content was reflected by an equivalent increase in 18:2 content indicating that the desaturation of 18:2 to 18:3 had been reduced. The relative content of the the other fatty acids remained unchanged.

**[0229]** Southern analysis for the presence of the intact, introduced antisense construction was performed, as described in Example 12 using Bam HI cut gDNA, on a number of the transformed lines listed below using groups of embryos from a single transformation event. The approximate intact antisense copy number was estimated from the number and intensity of hybridizing bands on the autoradiograms and is shown in Table 13.

TABLE 13

| Line No. | Antisense copy No. | 18:3 (low) | 18:3 (average) | 18:2/18:3 ratio |
|---|---|---|---|---|
| 2872 | 0 | 11.1 | 14.5 | 3.6 |
| 303-7/1 | 1 | 11.4 | 12.6 | 4.7 |
| 307-1/2 | 3 | 12.2 | 16.0 | 3.5 |
| 306-4/8 | 3 | 10.8 | 12.2 | 4.3 |
| 307-1/7 | 4 | 8.5 | 10.3 | 5.7 |
| 306-4/5 | 6 | 7.4 | 10.4 | 5.8 |
| 307-1/1 | 6 | 6.3 | 9.6 | 6.3 |
| 299-15/1 | 7 | 8.1 | 9.7 | 6.1 |
| 307-1/4 | 8 | 6.7 | 7.7 | 8.0 |

**[0230]** There was a reasonable correlation between intact antisense copy number and 18:3 content, an increase in copy number correlating with a decreased 18:3 content and a consequent increase in the 18:2/18:3 ratio. The average 18:2/18:3 ratio of line 307-1/4, which had at least 8 copies of the antisense cDNA, was more than twice that of the control.

SEQUENCE LISTING

**[0231]**

(1) GENERAL INFORMATION:

(i) APPLICANTS: Browae, John, Kinney, Anthony J., Pierce, John, Wierzbicki, Anna M., Yadav, Narendra S., Perez-Grau, Luis

(ii) TITLE OF INVENTION: Fatty Acid Desaturase Genes from Plants

(iii) NUMBER OF SEQUENCES: 32

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: E. I. du Pont de Nemours and Company
(B) STREET: 1007 Market Street
(C) CITY: Wilmington
(D) STATE: Delaware
(E) COUNTRY: U.S.A.
(F) ZIP: 19898

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: Macintosh
(C) OPERATING SYSTEM: Macintosh System, 6.0
(D) SOFTWARE: Microsoft Word, 4.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 07/804,259
(B) FILING DATE: 4 DECEMBER 1991

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Floyd, Linda A.
(B) REGISTRATION NUMBER: 33,692
(C) REFERENCE/DOCKET NUMBER: BB-1036-A

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (302) 992-4929
(B) TELEFAX: (302) 892-7949
(C) TELEX: 835420

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1350 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: CDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana IMMEDIATE SOURCE:
(B) CLONE: pCF3

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 46..1206

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
CTCTCTCTCT CTCTCTTCTC TCTTTCTCTC CCCCTCTCTC CGGCG ATG GTT GTT          54
                                                    Met Val Val
                                                      1

GCT ATG GAC CAA CGC ACC AAT GTG AAC GGA GAT CCC GGC GCC GGA GAC        102
Ala Met Asp Gln Arg Thr Asn Val Asn Gly Asp Pro Gly Ala Gly Asp
      5                  10                  15

CGG AAG AAA GAA GAA AGG TTT GAT CCG AGT GCA CAA CCA CCG TTC AAG        150
Arg Lys Lys Glu Glu Arg Phe Asp Pro Ser Ala Gln Pro Pro Phe Lys
 20                  25                  30                  35

ATC GGA GAT ATA AGG GCG GCG ATT CCT AAG CAC TGT TGG GTT AAG AGT        198
Ile Gly Asp Ile Arg Ala Ala Ile Pro Lys His Cys Trp Val Lys Ser
                 40                  45                  50

CCT TTG AGA TCA ATG AGT TAC GTC GTC AGA GAC ATT ATC GCC GTC GCG        246
Pro Leu Arg Ser Met Ser Tyr Val Val Arg Asp Ile Ile Ala Val Ala
             55                  60                  65

GCT TTG GCC ATC GCT GCC GTG TAT GTT GAT AGC TGG TTC CTT TGG CCT        294
Ala Leu Ala Ile Ala Ala Val Tyr Val Asp Ser Trp Phe Leu Trp Pro
         70                  75                  80

CTT TAT TGG GCC GCC CAA GGA ACA CTT TTC TGG GCC ATC TTT GTT CTC        342
Leu Tyr Trp Ala Ala Gln Gly Thr Leu Phe Trp Ala Ile Phe Val Leu
      85                  90                  95

GGC CAC GAC TGT GGA CAT GGG AGT TTC TCA GAC ATT CCT CTA CTG AAT        390
Gly His Asp Cys Gly His Gly Ser Phe Ser Asp Ile Pro Leu Leu Asn
100                 105                 110                 115
```

```
AGT GTG GTT GGT CAC ATT CTT CAT TCT TTC ATC CTC GTT CCT TAC CAT          438
Ser Val Val Gly His Ile Leu His Ser Phe Ile Leu Val Pro Tyr His
                120                 125                 130

GGT TGG AGA ATA AGC CAC CGG ACA CAC CAC CAG AAC CAT GGC CAT GTT          486
Gly Trp Arg Ile Ser His Arg Thr His His Gln Asn His Gly His Val
            135                 140                 145

GAA AAC GAC GAG TCA TGG GTT CCG TTA CCA GAA AGG GTG TAC AAG AAA          534
Glu Asn Asp Glu Ser Trp Val Pro Leu Pro Glu Arg Val Tyr Lys Lys
        150                 155                 160

TTG CCC CAC AGT ACT CGG ATG CTC AGA TAC ACT GTC CCT CTC CCC ATG          582
Leu Pro His Ser Thr Arg Met Leu Arg Tyr Thr Val Pro Leu Pro Met
    165                 170                 175

CTC GCA TAT CCT CTC TAT TTG TGC TAC AGA AGT CCT GGA AAA GAA GGA          630
Leu Ala Tyr Pro Leu Tyr Leu Cys Tyr Arg Ser Pro Gly Lys Glu Gly
180                 185                 190                 195

TCA CAT TTT AAC CCA TAC AGT AGT TTA TTT GCT CCA AGC GAG AGA AAG          678
Ser His Phe Asn Pro Tyr Ser Ser Leu Phe Ala Pro Ser Glu Arg Lys
                200                 205                 210

CTT ATT GCA ACT TCA ACT ACT TGT TGG TCC ATA ATG TTC GTC AGT CTT          726
Leu Ile Ala Thr Ser Thr Thr Cys Trp Ser Ile Met Phe Val Ser Leu
            215                 220                 225

ATC GCT CTA TCT TTC GTC TTC GGT CCA CTC GCG GTT CTT AAA GTC TAC          774
Ile Ala Leu Ser Phe Val Phe Gly Pro Leu Ala Val Leu Lys Val Tyr
        230                 235                 240

GGT GTA CCG TAC ATT ATC TTT GTG ATG TGG TTG GAT GCT GTC ACG TAT          822
Gly Val Pro Tyr Ile Ile Phe Val Met Trp Leu Asp Ala Val Thr Tyr
    245                 250                 255

TTG CAT CAT CAT GGT CAC GAT GAG AAG TTG CCT TGG TAT AGA GGC AAG          870
Leu His His His Gly His Asp Glu Lys Leu Pro Trp Tyr Arg Gly Lys
260                 265                 270                 275

GAA TGG AGT TAT CTA CGT GGA GGA TTA ACA ACA ATT GAT AGA GAT TAC          918
Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr Thr Ile Asp Arg Asp Tyr
                280                 285                 290

GGA ATC TTT AAC AAC ATT CAT CAC GAC ATT GGA ACT CAC GTG ATC CAT          966
Gly Ile Phe Asn Asn Ile His His Asp Ile Gly Thr His Val Ile His
            295                 300                 305

CAT CTC TTC CCA CAA ATC CCT CAC TAT CAC TTG GTC GAC GCC ACG AAA         1014
His Leu Phe Pro Gln Ile Pro His Tyr His Leu Val Asp Ala Thr Lys
        310                 315                 320

GCA GCT AAA CAT GTG TTG GGA AGA TAC TAC AGA GAA CCA AAG ACG TCA         1062
Ala Ala Lys His Val Leu Gly Arg Tyr Tyr Arg Glu Pro Lys Thr Ser
    325                 330                 335
```

```
GGA GCA ATA CCG ATC CAC TTG GTG GAG AGT TTG GTC GCA AGT ATT AAG          1110
Gly Ala Ile Pro Ile His Leu Val Glu Ser Leu Val Ala Ser Ile Lys
340                 345                 350                 355

AAA GAT CAT TAC GTC AGC GAC ACT GGT GAT ATT GTC TTC TAC GAG ACA          1158
Lys Asp His Tyr Val Ser Asp Thr Gly Asp Ile Val Phe Tyr Glu Thr
                360                 365                 370

GAT CCA GAT CTC TAC GTT TAC GCT TCT GAC AAA TCT AAA ATC AAT TAATCTCCAT 1213
Asp Pro Asp Leu Tyr Val Tyr Ala Ser Asp Lys Ser Lys Ile Asn
            375                 380                 385

TTGTTTAGCT CTATTAGGAA TAAACCAGCC CACTTTTAAA ATTTTTATTT CTTGTTGTTT      1273

TTAAGTTAAA AGTGTACTCG TGAAACTCTT TTTTTTTTCT TTTTTTTTAT TAATGTATTT      1333

ACATTACAAG GCGTAAA                                                     1350
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 386 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Val Val Ala Met Asp Gln Arg Thr Asn Val Asn Gly Asp Pro Gly
 1               5                  10                  15

Ala Gly Asp Arg Lys Lys Glu Glu Arg Phe Asp Pro Ser Ala Gln Pro
            20                  25                  30

Pro Phe Lys Ile Gly Asp Ile Arg Ala Ala Ile Pro Lys His Cys Trp
        35                  40                  45

Val Lys Ser Pro Leu Arg Ser Met Ser Tyr Val Val Arg Asp Ile Ile
    50                  55                  60

Ala Val Ala Ala Leu Ala Ile Ala Ala Val Tyr Val Asp Ser Trp Phe
65                  70                  75                  80

Leu Trp Pro Leu Tyr Trp Ala Ala Gln Gly Thr Leu Phe Trp Ala Ile
                85                  90                  95

Phe Val Leu Gly His Asp Cys Gly His Gly Ser Phe Ser Asp Ile Pro
            100                 105                 110

Leu Leu Asn Ser Val Val Gly His Ile Leu His Ser Phe Ile Leu Val
        115                 120                 125

Pro Tyr His Gly Trp Arg Ile Ser His Arg Thr His His Gln Asn His
    130                 135                 140
```

```
Gly His Val Glu Asn Asp Glu Ser Trp Val Pro Leu Pro Glu Arg Val
145                 150                 155                 160

Tyr Lys Lys Leu Pro His Ser Thr Arg Met Leu Arg Tyr Thr Val Pro
                165                 170                 175

Leu Pro Met Leu Ala Tyr Pro Leu Tyr Leu Cys Tyr Arg Ser Pro Gly
            180                 185                 190

Lys Glu Gly Ser His Phe Asn Pro Tyr Ser Ser Leu Phe Ala Pro Ser
        195                 200                 205

Glu Arg Lys Leu Ile Ala Thr Ser Thr Thr Cys Trp Ser Ile Met Phe
    210                 215                 220

Val Ser Leu Ile Ala Leu Ser Phe Val Phe Gly Pro Leu Ala Val Leu
225                 230                 235                 240

Lys Val Tyr Gly Val Pro Tyr Ile Ile Phe Val Met Trp Leu Asp Ala
                245                 250                 255

Val Thr Tyr Leu His His His Gly His Asp Glu Lys Leu Pro Trp Tyr
            260                 265                 270

Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr Thr Ile Asp
        275                 280                 285

Arg Asp Tyr Gly Ile Phe Asn Asn Ile His His Asp Ile Gly Thr His
    290                 295                 300

Val Ile His His Leu Phe Pro Gln Ile Pro His Tyr His Leu Val Asp
305                 310                 315                 320

Ala Thr Lys Ala Ala Lys His Val Leu Gly Arg Tyr Tyr Arg Glu Pro
                325                 330                 335

Lys Thr Ser Gly Ala Ile Pro Ile His Leu Val Glu Ser Leu Val Ala
            340                 345                 350

Ser Ile Lys Lys Asp His Tyr Val Ser Asp Thr Gly Asp Ile Val Phe
        355                 360                 365

Tyr Glu Thr Asp Pro Asp Leu Tyr Val Tyr Ala Ser Asp Lys Ser Lys
    370                 375                 380

Ile Asn
385
```

(2) INFORMATION FOR SEQ ID NO:3:.

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 255 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana

(vii) IMMEDIATE SOURCE:

(B) CLONE: pF1

(ix) FEATURE:

(A) NAME/KEY: exon
(B) LOCATION: 68..255

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
AAATTCATCA AACCCTTTCT TCACCACATT ATTTTCACTG AGCGCATAAC ATTTTTGAGA      60
CAAGAGACTC TCTCTCTCTC TCTTCTCTCT TTCTCTCCCC CTCTCTCCGG CGATGGTTGT     120
TGCTATGGAC CAACGCACCA ATGTGAACGG AGATCCCGGC GCCGGAGACC GGAAGAAAGA     180
AGAAAGGTTT GATCCGAGTG CACAACCACC GTTCAAGATC GGAGATATAA GGGCGGCGAT     240
TCCTAAGCAC TGTTG                                                      255
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1525 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana

(vii) IMMEDIATE SOURCE:

(B) CLONE: pACF2-2

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 10..1350

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
CAAGTTCTA ATG GCG AAC TTG GTC TTA TCA GAA TGT GGT ATA CGA CCT          48
          Met Ala Asn Leu Val Leu Ser Glu Cys Gly Ile Arg Pro
            1               5                  10

CTC CCC AGA ATC TAC ACA ACA CCC AGA TCC AAT TTC CTC TCC AAC AAC         96
Leu Pro Arg Ile Tyr Thr Thr Pro Arg Ser Asn Phe Leu Ser Asn Asn
     15                  20                  25

AAC AAA TTC AGA CCA TCA CTT TCT TCT TCT TCT TAC AAA ACA TCA TCA        144
Asn Lys Phe Arg Pro Ser Leu Ser Ser Ser Ser Tyr Lys Thr Ser Ser
 30                  35                  40                  45

TCT CCT CTG TCT TTT GGT CTG AAT TCA CGA GAT GGG TTC ACG AGG AAT        192
Ser Pro Leu Ser Phe Gly Leu Asn Ser Arg Asp Gly Phe Thr Arg Asn
                 50                  55                  60

TGG GCG TTG AAT GTG AGC ACA CCA TTA ACG ACA CCA ATA TTT GAG GAG        240
Trp Ala Leu Asn Val Ser Thr Pro Leu Thr Thr Pro Ile Phe Glu Glu
               65                  70                  75

TCT CCA TTG GAG GAA GAT AAT AAA CAG AGA TTC GAT CCA GGT GCG CCT        288
Ser Pro Leu Glu Glu Asp Asn Lys Gln Arg Phe Asp Pro Gly Ala Pro
         80                  85                  90

CCT CCG TTC AAT TTA GCT GAT ATT AGA GCA GCT ATA CCT AAG CAT TGT        336
Pro Pro Phe Asn Leu Ala Asp Ile Arg Ala Ala Ile Pro Lys His Cys
     95                  100                 105

TGG GTT AAG AAT CCA TGG AAG TCT TTG AGT TAT GTC GTC AGA GAC GTC        384
Trp Val Lys Asn Pro Trp Lys Ser Leu Ser Tyr Val Val Arg Asp Val
110                 115                 120                 125

GCT ATC GTC TTT GCA TTG GCT GCT GGA GCT GCT TAC CTC AAC AAT TGG        432
Ala Ile Val Phe Ala Leu Ala Ala Gly Ala Ala Tyr Leu Asn Asn Trp
                130                 135                 140

ATT GTT TGG CCT CTC TAT TGG CTC GCT CAA GGA ACC ATG TTT TGG GCT        480
Ile Val Trp Pro Leu Tyr Trp Leu Ala Gln Gly Thr Met Phe Trp Ala
            145                 150                 155

CTC TTT GTT CTT GGT CAT GAC TGT GGA CAT GGT AGT TTC TCA AAT GAT        528
Leu Phe Val Leu Gly His Asp Cys Gly His Gly Ser Phe Ser Asn Asp
        160                 165                 170

CCG AAG TTG AAC AGT GTG GTC GGT CAT CTT CTT CAT TCC TCA ATT CTG        576
Pro Lys Leu Asn Ser Val Val Gly His Leu Leu His Ser Ser Ile Leu
    175                 180                 185

GTC CCA TAC CAT GGC TGG AGA ATT AGT CAC AGA ACT CAC CAC CAG AAC        624
Val Pro Tyr His Gly Trp Arg Ile Ser His Arg Thr His His Gln Asn
190                 195                 200                 205

CAT GGA CAT GTT GAG AAT GAC GAA TCT TGG CAT CCT ATG TCT GAG AAA        672
His Gly His Val Glu Asn Asp Glu Ser Trp His Pro Met Ser Glu Lys
                210                 215                 220
```

```
ATC TAC AAT ACT TTG GAC AAG CCG ACT AGA TTC TTT AGA TTT ACA CTG          720
Ile Tyr Asn Thr Leu Asp Lys Pro Thr Arg Phe Phe Arg Phe Thr Leu
            225                 230                 235

CCT CTC GTG ATG CTT GCA TAC CCT TTC TAC TTG TGG GCT CGA AGT CCG          768
Pro Leu Val Met Leu Ala Tyr Pro Phe Tyr Leu Trp Ala Arg Ser Pro
        240                 245                 250

GGG AAA AAG GGT TCT CAT TAC CAT CCA GAC AGT GAC TTG TTC CTC CCT          816
Gly Lys Lys Gly Ser His Tyr His Pro Asp Ser Asp Leu Phe Leu Pro
    255                 260                 265

AAA GAG AGA AAG GAT GTC CTC ACT TCT ACT GCT TGT TGG ACT GCA ATG          864
Lys Glu Arg Lys Asp Val Leu Thr Ser Thr Ala Cys Trp Thr Ala Met
270                 275                 280                 285

GCT GCT CTG CTT GTT TGT CTC AAC TTC ACA ATC GGT CCA ATT CAA ATG          912
Ala Ala Leu Leu Val Cys Leu Asn Phe Thr Ile Gly Pro Ile Gln Met
                290                 295                 300

CTC AAA CTT TAT GGA ATT CCT TAC TGG ATA AAT GTA ATG TGG TTG GAC          960
Leu Lys Leu Tyr Gly Ile Pro Tyr Trp Ile Asn Val Met Trp Leu Asp
            305                 310                 315

TTT GTG ACT TAC CTG CAT CAC CAT GGT CAT GAA GAT AAG CTT CCT TGG         1008
Phe Val Thr Tyr Leu His His His Gly His Glu Asp Lys Leu Pro Trp
        320                 325                 330

TAC CGT GGC AAG GAG TGG AGT TAC CTG AGA GGA GGA CTT ACA ACA TTG         1056
Tyr Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr Thr Leu
    335                 340                 345

GAT CGT GAC TAC GGA TTG ATC AAT AAC ATC CAT CAT GAT ATT GGA ACT         1104
Asp Arg Asp Tyr Gly Leu Ile Asn Asn Ile His His Asp Ile Gly Thr
350                 355                 360                 365

CAT GTG ATA CAT CAT CTT TTC CCG CAG ATC CCA CAT TAT CAT CTA GTA         1152
His Val Ile His His Leu Phe Pro Gln Ile Pro His Tyr His Leu Val
                370                 375                 380

GAA GCA ACA GAA GCA GCT AAA CCA GTA TTA GGG AAG TAT TAC AGG GAG         1200
Glu Ala Thr Glu Ala Ala Lys Pro Val Leu Gly Lys Tyr Tyr Arg Glu
            385                 390                 395

CCT GAT AAG TCT GGA CCG TTG CCA TTA CAT TTA CTG GAA ATT CTA GCG         1248
Pro Asp Lys Ser Gly Pro Leu Pro Leu His Leu Leu Glu Ile Leu Al
        400                 405                 410

AAA AGT ATA AAA GAA GAT CAT TAC GTG AGC GAC GAA GGA GAA GTT GTA         1296
Lys Ser Ile Lys Glu Asp His Tyr Val Ser Asp Glu Gly Glu Val Val
    415                 420                 425

TAC TAT AAA GCA GAT CCA AAT CTC TAT GGA GAG GTC AAA GTA AGA GCA         1344
Tyr Tyr Lys Ala Asp Pro Asn Leu Tyr Gly Glu Val Lys Val Arg Ala
430                 435                 440                 445

GAT TGAAATGAAG CAGGCTTGAG ATTGAAGTTT TTTCTATTTC AGACCAGCTG              1397
Asp
```

```
ATTTTTTGCT TACTGTATCA ATTTATTGTG TCACCCACCA GAGAGTTAGT ATCTCTGAAT    1457

ACGATCGATC AGATGGAAAC AACAAATTTG TTTGCGATAC TGAAGCTATA TATACCATAC    1517

ATTGCATT                                                             1525
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 446 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Met Ala Asn Leu Val Leu Ser Glu Cys Gly Ile Arg Pro Leu Pro Arg
 1               5                  10                  15

Ile Tyr Thr Thr Pro Arg Ser Asn Phe Leu Ser Asn Asn Asn Lys Phe
            20                  25                  30

Arg Pro Ser Leu Ser Ser Ser Ser Tyr Lys Thr Ser Ser Ser Pro Leu
        35                  40                  45

Ser Phe Gly Leu Asn Ser Arg Asp Gly Phe Thr Arg Asn Trp Ala Leu
    50                  55                  60

Asn Val Ser Thr Pro Leu Thr Thr Pro Ile Phe Glu Glu Ser Pro Leu
65                  70                  75                  80

Glu Glu Asp Asn Lys Gln Arg Phe Asp Pro Gly Ala Pro Pro Pro Phe
                85                  90                  95

Asn Leu Ala Asp Ile Arg Ala Ala Ile Pro Lys His Cys Trp Val Lys
            100                 105                 110

Asn Pro Trp Lys Ser Leu Ser Tyr Val Val Arg Asp Val Ala Ile Val
        115                 120                 125

Phe Ala Leu Ala Ala Gly Ala Ala Tyr Leu Asn Asn Trp Ile Val Trp
    130                 135                 140

Pro Leu Tyr Trp Leu Ala Gln Gly Thr Met Phe Trp Ala Leu Phe Val
145                 150                 155                 160

Leu Gly His Asp Cys Gly His Gly Ser Phe Ser Asn Asp Pro Lys Leu
                165                 170                 175

Asn Ser Val Val Gly His Leu Leu His Ser Ser Ile Leu Val Pro Tyr
            180                 185                 190

His Gly Trp Arg Ile Ser His Arg Thr His His Gln Asn His Gly His
        195                 200                 205
```

```
Val Glu Asn Asp Glu Ser Trp His Pro Met Ser Glu Lys Ile Tyr Asn
    210                 215                 220

Thr Leu Asp Lys Pro Thr Arg Phe Phe Arg Phe Thr Leu Pro Leu Val
225                 230                 235                 240

Met Leu Ala Tyr Pro Phe Tyr Leu Trp Ala Arg Ser Pro Gly Lys Lys
                245                 250                 255

Gly Ser His Tyr His Pro Asp Ser Asp Leu Phe Leu Pro Lys Glu Arg
            260                 265                 270

Lys Asp Val Leu Thr Ser Thr Ala Cys Trp Thr Ala Met Ala Ala Leu
        275                 280                 285

Leu Val Cys Leu Asn Phe Thr Ile Gly Pro Ile Gln Met Leu Lys Leu
    290                 295                 300

Tyr Gly Ile Pro Tyr Trp Ile Asn Val Met Trp Leu Asp Phe Val Thr
305                 310                 315                 320

Tyr Leu His His His Gly His Glu Asp Lys Leu Pro Trp Tyr Arg Gly
                325                 330                 335

Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr Thr Leu Asp Arg Asp
            340                 345                 350

Tyr Gly Leu Ile Asn Asn Ile His His Asp Ile Gly Thr His Val Ile
        355                 360                 365

His His Leu Phe Pro Gln Ile Pro His Tyr His Leu Val Glu Ala Thr
    370                 375                 380

Glu Ala Ala Lys Pro Val Leu Gly Lys Tyr Tyr Arg Glu Pro Asp Lys
385                 390                 395                 400

Ser Gly Pro Leu Pro Leu His Leu Leu Glu Ile Leu Ala Lys Ser Ile
                405                 410                 415

Lys Glu Asp His Tyr Val Ser Asp Glu Gly Glu Val Val Tyr Tyr Lys
            420                 425                 430

Ala Asp Pro Asn Leu Tyr Gly Glu Val Lys Val Arg Ala Asp
        435                 440                 445
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1429 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Brassica napus

(vii) IMMEDIATE SOURCE:

(B) CLONE: pBNSF3-f2

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 79..1212

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
TTCAAATTCA GACAATCCCC TTCTTCTCCC CGGTTTCGTC TGAACTCTCG AAACTGGGCG        60

TTGAATGTAA CCACACCT CTA ACA GTC GAC TCC TCA TCA TCT CCT CCA ATC          111
                     Leu Thr Val Asp Ser Ser Ser Ser Pro Pro Ile
                       1               5                  10

GAG GAA GAA CCC AAA ACG CAG AGA TTC GAC CCA GGC GCT CCT CCT CCG          159
Glu Glu Glu Pro Lys Thr Gln Arg Phe Asp Pro Gly Ala Pro Pro Pro
              15                  20                  25

TTC AAC CTA GCT GAC ATC AGA GCG GCG ATA CCT AAG CAT TGC TGG GTT          207
Phe Asn Leu Ala Asp Ile Arg Ala Ala Ile Pro Lys His Cys Trp Val
         30                  35                  40

AAG AAT CCA TGG AAG TCT ATG AGT TAC GTC GTC AGA GAG CTA GCC ATC          255
Lys Asn Pro Trp Lys Ser Met Ser Tyr Val Val Arg Glu Leu Ala Ile
     45                  50                  55

GTG TTC GCA CTA GCT GCT GGA GCT GCT TAC CTC AAC AAT TGG CTT GTT          303
Val Phe Ala Leu Ala Ala Gly Ala Ala Tyr Leu Asn Asn Trp Leu Val
 60                  65                  70                  75

TGG CCT CTC TAT TGG ATT GCT CAA GGA ACC ATG TTC TGG GCT CTC TTT          351
Trp Pro Leu Tyr Trp Ile Ala Gln Gly Thr Met Phe Trp Ala Leu Phe
                 80                  85                  90

GTT CTT GGC CAT GAC TGT GGA CAT GGA AGC TTC TCA AAT GAT CCG AGG          399
Val Leu Gly His Asp Cys Gly His Gly Ser Phe Ser Asn Asp Pro Arg
             95                  100                 105

TTG AAC AGT GTG GTG GGT CAC CTT CTT CAT TCC TCT ATT CTA GTC CCT          447
Leu Asn Ser Val Val Gly His Leu Leu His Ser Ser Ile Leu Val Pro
        110                 115                 120

TAC CAT GGC TGG AGA ATT AGC CAC AGA ACT CAC CAC CAG AAC CAT GGA          495
Tyr His Gly Trp Arg Ile Ser His Arg Thr His His Gln Asn His Gly
    125                 130                 135

CAT GTT GAG AAC GAT GAA TCT TGG CAT CCT ATG TCT GAG AAA ATC TAC          543
His Val Glu Asn Asp Glu Ser Trp His Pro Met Ser Glu Lys Ile Tyr
140                 145                 150                 155
```

```
AAG AGT TTG GAC AAA CCC ACT CGG TTC TTT AGA TTT ACA TTG CCT CTC          591
Lys Ser Leu Asp Lys Pro Thr Arg Phe Phe Arg Phe Thr Leu Pro Leu
                160                 165                 170

GTG ATG CTC GCT TAC CCT TTC TAC TTG TGG GCA AGA AGT CCA GGG AAG          639
Val Met Leu Ala Tyr Pro Phe Tyr Leu Trp Ala Arg Ser Pro Gly Lys
            175                 180                 185

AAG GGT TCT CAT TAC CAT CCA GAC AGC GAC TTG TTC CTT CCT AAA GAG          687
Lys Gly Ser His Tyr His Pro Asp Ser Asp Leu Phe Leu Pro Lys Glu
        190                 195                 200

AGA AAC GAT GTT CTC ACT TCT ACC GCT TGT TGG ACT GCA ATG GCT GTT          735
Arg Asn Asp Val Leu Thr Ser Thr Ala Cys Trp Thr Ala Met Ala Val
    205                 210                 215

CTG CTT GTC TGT CTC AAC TTC GTG ATG GGT CCA ATG CAA ATG CTC AAA          783
Leu Leu Val Cys Leu Asn Phe Val Met Gly Pro Met Gln Met Leu Lys
220                 225                 230                 235

CTT TAT GTC ATT CCT TAC TGG ATA AAT GTA ATG TGG TTG GAC TTT GTG          831
Leu Tyr Val Ile Pro Tyr Trp Ile Asn Val Met Trp Leu Asp Phe Val
                240                 245                 250

ACT TAC CTG CAT CAC CAT GGT CAT GAA GAT AAG CTC CCT TGG TAC CGT          879
Thr Tyr Leu His His His Gly His Glu Asp Lys Leu Pro Trp Tyr Arg
            255                 260                 265

GGG AAG GAA TGG AGT TAC TTG AGA GGA GGA CTT ACA ACA TTG GAC CGG          927
Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr Thr Leu Asp Arg
        270                 275                 280

GAC TAC GGA TTG ATC AAC AAC ATC CAT CAC GAC ATT GGA ACT CAT GTG          975
Asp Tyr Gly Leu Ile Asn Asn Ile His His Asp Ile Gly Thr His Val
    285                 290                 295

ATA CAT CAT CTT TTC CCT CAG ATC CCA CAT TAT CAT CTA GTA GAA GCA         1023
Ile His His Leu Phe Pro Gln Ile Pro His Tyr His Leu Val Glu Ala
300                 305                 310                 315

ACA GAA GCA GCT AAA CCA GTA TTA GGG AAG TAT TAT AGG GAG CCT GAT         1071
Thr Glu Ala Ala Lys Pro Val Leu Gly Lys Tyr Tyr Arg Glu Pro Asp
                320                 325                 330

AAG TCT GGA CCT TTG CCA TTA CAT TTA CTG GGA ATC TTA GCA AAA AGT         1119
Lys Ser Gly Pro Leu Pro Leu His Leu Leu Gly Ile Leu Ala Lys Ser
            335                 340                 345

ATT AAA GAA GAT CAT TTT GTG AGC GAT GAA GGA GAT GTT GTA TAC TAT         1167
Ile Lys Glu Asp His Phe Val Ser Asp Glu Gly Asp Val Val Tyr Tyr
        350                 355                 360

GAA GCA GAC CCT AAT CTC TAT GGA GAG ATC AAG GTA ACA GCA GAG             1212
Glu Ala Asp Pro Asn Leu Tyr Gly Glu Ile Lys Val Thr Ala Glu
    365                 370                 375

TGAAATGAAG CTGTCAGATT TATCTATTTC TGACCAGCTG ATTTTTTTTG CTTATTAATG       1272
```

```
TCAATTCATT GTGTTACCAT TATCTCTGAA TACAATCAGA TGGAAACCCC AACTTTGTTT      1332

TCAATACTTG AAGCTATATA TATATATATA TATGTAAGAT ACATTGTATT GTCATTAGAT      1392

TCACCATTCT CAAGGTTCTT ATACAAAAAA AAAAAAA                               1429
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 378 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
Leu Thr Val Asp Ser Ser Ser Ser Pro Pro Ile Glu Glu Glu Pro Lys
 1               5                  10                  15

Thr Gln Arg Phe Asp Pro Gly Ala Pro Pro Pro Phe Asn Leu Ala Asp
            20                  25                  30

Ile Arg Ala Ala Ile Pro Lys His Cys Trp Val Lys Asn Pro Trp Lys
        35                  40                  45

Ser Met Ser Tyr Val Val Arg Glu Leu Ala Ile Val Phe Ala Leu Ala
    50                  55                  60

Ala Gly Ala Ala Tyr Leu Asn Asn Trp Leu Val Trp Pro Leu Tyr Trp
65                  70                  75                  80

Ile Ala Gln Gly Thr Met Phe Trp Ala Leu Phe Val Leu Gly His Asp
                85                  90                  95

Cys Gly His Gly Ser Phe Ser Asn Asp Pro Arg Leu Asn Ser Val Val
            100                 105                 110

Gly His Leu Leu His Ser Ser Ile Leu Val Pro Tyr His Gly Trp Arg
        115                 120                 125

Ile Ser His Arg Thr His His Gln Asn His Gly His Val Glu Asn Asp
    130                 135                 140

Glu Ser Trp His Pro Met Ser Glu Lys Ile Tyr Lys Ser Leu Asp Lys
145                 150                 155                 160

Pro Thr Arg Phe Phe Arg Phe Thr Leu Pro Leu Val Met Leu Ala Tyr
                165                 170                 175

Pro Phe Tyr Leu Trp Ala Arg Ser Pro Gly Lys Lys Gly Ser His Tyr
            180                 185                 190

His Pro Asp Ser Asp Leu Phe Leu Pro Lys Glu Arg Asn Asp Val Leu
        195                 200                 205
```

```
Thr Ser Thr Ala Cys Trp Thr Ala Met Ala Val Leu Leu Val Cys Leu
    210                 215                 220

Asn Phe Val Met Gly Pro Met Gln Met Leu Lys Leu Tyr Val Ile Pro
225                 230                 235                 240

Tyr Trp Ile Asn Val Met Trp Leu Asp Phe Val Thr Tyr Leu His His
                245                 250                 255

His Gly His Glu Asp Lys Leu Pro Trp Tyr Arg Gly Lys Glu Trp Ser
            260                 265                 270

Tyr Leu Arg Gly Gly Leu Thr Thr Leu Asp Arg Asp Tyr Gly Leu Ile
        275                 280                 285

Asn Asn Ile His His Asp Ile Gly Thr His Val Ile His His Leu Phe
    290                 295                 300

Pro Gln Ile Pro His Tyr His Leu Val Glu Ala Thr Glu Ala Ala Lys
305                 310                 315                 320

Pro Val Leu Gly Lys Tyr Tyr Arg Glu Pro Asp Lys Ser Gly Pro Leu
                325                 330                 335

Pro Leu His Leu Leu Gly Ile Leu Ala Lys Ser Ile Lys Glu Asp His
            340                 345                 350

Phe Val Ser Asp Glu Gly Asp Val Val Tyr Tyr Glu Ala Asp Pro Asn
        355                 360                 365

Leu Tyr Gly Glu Ile Lys Val Thr Ala Glu
    370                 375
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1429 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Brassica napus

(vii) IMMEDIATE SOURCE:

(B) CLONE: pBNSFd-2

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..1215

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
TTC AAA TTC AGA CAA TCC CCT TCT TCT CCC CGG TTT CGT CTG AAC TCT          48
Phe Lys Phe Arg Gln Ser Pro Ser Ser Pro Arg Phe Arg Leu Asn Ser
 1               5                  10                  15

CGA AAC TGG GCG TTG AAT GTA ACC ACA CCT CTA ACA GTC GAC TCC TCA          96
Arg Asn Trp Ala Leu Asn Val Thr Thr Pro Leu Thr Val Asp Ser Ser
            20                  25                  30

TCA TCT CCT CCA ATC GAG GAA GAA CCC AAA ACG CAG AGA TTC GAC CCA         144
Ser Ser Pro Pro Ile Glu Glu Glu Pro Lys Thr Gln Arg Phe Asp Pro
        35                  40                  45

GGC GCT CCT CCT CCG TTC AAC CTA GCT GAC ATC AGA GCG GCG ATA CCT         192
Gly Ala Pro Pro Pro Phe Asn Leu Ala Asp Ile Arg Ala Ala Ile Pro
    50                  55                  60

AAG CAT TGC TGG GTT AAG AAT CCA TGG AAG TCT ATG AGT TAC GTC GTC         240
Lys His Cys Trp Val Lys Asn Pro Trp Lys Ser Met Ser Tyr Val Val
65                  70                  75                  80

AGA GAG CTA GCC ATC GTG TTC GCA CTA GCT GCT GGA GCT GCT TAC CTC         288
Arg Glu Leu Ala Ile Val Phe Ala Leu Ala Ala Gly Ala Ala Tyr Leu
                85                  90                  95

AAC AAT TGG CTT GTT TGG CCT CTC TAT TGG ATT GCT CAA GGA ACC ATG         336
Asn Asn Trp Leu Val Trp Pro Leu Tyr Trp Ile Ala Gln Gly Thr Met
            100                 105                 110

TTC TGG GCT CTC TTT GTT CTT GGC CAT GAC TGT GGA CAT GGA AGC TTC         384
Phe Trp Ala Leu Phe Val Leu Gly His Asp Cys Gly His Gly Ser Phe
        115                 120                 125

TCA AAT GAT CCG AGG TTG AAC AGT GTG GTG GGT CAC CTT CTT CAT TCC         432
Ser Asn Asp Pro Arg Leu Asn Ser Val Val Gly His Leu Leu His Ser
    130                 135                 140

TCT ATT CTA GTC CCT TAC CAT GGC TGG AGA ATT AGC CAC AGA ACT CAC         480
Ser Ile Leu Val Pro Tyr His Gly Trp Arg Ile Ser His Arg Thr His
145                 150                 155                 160

CAC CAG AAC CAT GGA CAT GTT GAG AAC GAT GAA TCT TGG CAT CCT ATG         528
His Gln Asn His Gly His Val Glu Asn Asp Glu Ser Trp His Pro Met
                165                 170                 175

TCT GAG AAA ATC TAC AAG AGT TTG GAC AAA CCC ACT CGG TTC TTT AGA         576
Ser Glu Lys Ile Tyr Lys Ser Leu Asp Lys Pro Thr Arg Phe Phe Arg
            180                 185                 190

TTT ACA TTG CCT CTC GTG ATG CTC GCT TAC CCT TTC TAC TTG TGG GCA         624
Phe Thr Leu Pro Leu Val Met Leu Ala Tyr Pro Phe Tyr Leu Trp Ala
        195                 200                 205
```

```
AGA AGT CCA GGG AAG AAG GGT TCT CAT TAC CAT CCA GAC AGC GAC TTG          672
Arg Ser Pro Gly Lys Lys Gly Ser His Tyr His Pro Asp Ser Asp Leu
    210                 215                 220

TTC CTT CCT AAA GAG AGA AAC GAT GTT CTC ACT TCT ACC GCT TGT TGG          720
Phe Leu Pro Lys Glu Arg Asn Asp Val Leu Thr Ser Thr Ala Cys Trp
225                 230                 235                 240

ACT GCA ATG GCT GTT CTG CTT GTC TGT CTC AAC TTC GTG ATG GGT CCA          768
Thr Ala Met Ala Val Leu Leu Val Cys Leu Asn Phe Val Met Gly Pro
                245                 250                 255

ATG CAA ATG CTC AAA CTT TAT GTC ATT CCT TAC TGG ATA AAT GTA ATG          816
Met Gln Met Leu Lys Leu Tyr Val Ile Pro Tyr Trp Ile Asn Val Met
            260                 265                 270

TGG TTG GAC TTT GTG ACT TAC CTG CAT CAC CAT GGT CAT GAA GAT AAG          864
Trp Leu Asp Phe Val Thr Tyr Leu His His His Gly His Glu Asp Lys
        275                 280                 285

CTC CCT TGG TAC CGT GGG AAG GAA TGG AGT TAC TTG AGA GGA GGA CTT          912
Leu Pro Trp Tyr Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu
    290                 295                 300

ACA ACA TTG GAC CGG GAC TAC GGA TTG ATC AAC AAC ATC CAT CAC GAC          960
Thr Thr Leu Asp Arg Asp Tyr Gly Leu Ile Asn Asn Ile His His Asp
305                 310                 315                 320

ATT GGA ACT CAT GTG ATA CAT CAT CTT TTC CCT CAG ATC CCA CAT TAT         1008
Ile Gly Thr His Val Ile His His Leu Phe Pro Gln Ile Pro His Tyr
                325                 330                 335

CAT CTA GTA GAA GCA ACA GAA GCA GCT AAA CCA GTA TTA GGG AAG TAT         1056
His Leu Val Glu Ala Thr Glu Ala Ala Lys Pro Val Leu Gly Lys Tyr
            340                 345                 350

TAT AGG GAG CCT GAT AAG TCT GGA CCT TTG CCA TTA CAT TTA CTG GGA         1104
Tyr Arg Glu Pro Asp Lys Ser Gly Pro Leu Pro Leu His Leu Leu Gly
        355                 360                 365

ATC TTA GCA AAA AGT ATT AAA GAA GAT CAT TTT GTG AGC GAT GAA GGA         1152
Ile Leu Ala Lys Ser Ile Lys Glu Asp His Phe Val Ser Asp Glu Gly
    370                 375                 380

GAT GTT GTA TAC TAT GAA GCA GAC CCT AAT CTC TAT GGA GAG ATC AAG         1200
Asp Val Val Tyr Tyr Glu Ala Asp Pro Asn Leu Tyr Gly Glu Ile Lys
385                 390                 395                 400

GTA ACA GCA GAG TGAAATGAAG CTGTCAGATT TATCTATTTC TGACCAGCTG             1252
Val Thr Ala Glu
                405

ATTTTTTTTG CTTATTAATG TCAATTCATT GTGTTACCAT TATCTCTGAA TACAATCAGA       1312

TGGAAACCCC AACTTTGTTT TCAATACTTG AAGCTATATA TATATATATA TATGTAAGAT       1372

ACATTGTATT GTCATTAGAT TCACCATTCT CAAGGTTCTT ATACAAAAAA AAAAAAA          1429
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 404 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
Phe Lys Phe Arg Gln Ser Pro Ser Ser Pro Arg Phe Arg Leu Asn Ser
 1               5                  10                  15

Arg Asn Trp Ala Leu Asn Val Thr Thr Pro Leu Thr Val Asp Ser Ser
            20                  25                  30

Ser Ser Pro Pro Ile Glu Glu Glu Pro Lys Thr Gln Arg Phe Asp Pro
        35                  40                  45

Gly Ala Pro Pro Pro Phe Asn Leu Ala Asp Ile Arg Ala Ala Ile Pro
    50                  55                  60

Lys His Cys Trp Val Lys Asn Pro Trp Lys Ser Met Ser Tyr Val Val
65                  70                  75                  80

Arg Glu Leu Ala Ile Val Phe Ala Leu Ala Ala Gly Ala Ala Tyr Leu
                85                  90                  95

Asn Asn Trp Leu Val Trp Pro Leu Tyr Trp Ile Ala Gln Gly Thr Met
            100                 105                 110

Phe Trp Ala Leu Phe Val Leu Gly His Asp Cys Gly His Gly Ser Phe
        115                 120                 125

Ser Asn Asp Pro Arg Leu Asn Ser Val Val Gly His Leu Leu His Ser
    130                 135                 140

Ser Ile Leu Val Pro Tyr His Gly Trp Arg Ile Ser His Arg Thr His
145                 150                 155                 160

His Gln Asn His Gly His Val Glu Asn Asp Glu Ser Trp His Pro Met
                165                 170                 175

Ser Glu Lys Ile Tyr Lys Ser Leu Asp Lys Pro Thr Arg Phe Phe Arg
            180                 185                 190

Phe Thr Leu Pro Leu Val Met Leu Ala Tyr Pro Phe Tyr Leu Trp Ala
        195                 200                 205

Arg Ser Pro Gly Lys Lys Gly Ser His Tyr His Pro Asp Ser Asp Leu
    210                 215                 220

Phe Leu Pro Lys Glu Arg Asn Asp Val Leu Thr Ser Thr Ala Cys Trp
225                 230                 235                 240
```

```
Thr Ala Met Ala Val Leu Leu Val Cys Leu Asn Phe Val Met Gly Pro
                245                 250                 255

Met Gln Met Leu Lys Leu Tyr Val Ile Pro Tyr Trp Ile Asn Val Met
            260                 265                 270

Trp Leu Asp Phe Val Thr Tyr Leu His His His Gly His Glu Asp Lys
        275                 280                 285

Leu Pro Trp Tyr Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu
    290                 295                 300

Thr Thr Leu Asp Arg Asp Tyr Gly Leu Ile Asn Asn Ile His His Asp
305                 310                 315                 320

Ile Gly Thr His Val Ile His His Leu Phe Pro Gln Ile Pro His Tyr
                325                 330                 335

His Leu Val Glu Ala Thr Glu Ala Ala Lys Pro Val Leu Gly Lys Tyr
            340                 345                 350

Tyr Arg Glu Pro Asp Lys Ser Gly Pro Leu Pro Leu His Leu Leu Gly
        355                 360                 365

Ile Leu Ala Lys Ser Ile Lys Glu Asp His Phe Val Ser Asp Glu Gly
    370                 375                 380

Asp Val Val Tyr Tyr Glu Ala Asp Pro Asn Leu Tyr Gly Glu Ile Lys
385                 390                 395                 400

Val Thr Ala Glu
```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2181 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Glycine max

(vii) IMMEDIATE SOURCE:

(B) CLONE: pXF1

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 855..1997

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
ACAATAATAA ATCCATATTT TTATAATTAA AAGTAGTAGA TTACAGCGAT GCACTTGAGA      60

AACATATTAA GTGGACTAAT TCTCCCTGGT CAAGCAAGAA AAAAACCAGC TATGACCCAA     120

GGTAGAGAGA GATTATACAC AGAATACTAG TAATTAACTA AGACTGGCTC TGCAATTGCC     180

AAAAACTCCA TTGCAGTAGC AGCCACCTGA GAAGACACTA AGACCTAGAC TAGACCATAC     240

ATATGAAGAT TAATACGCTT ACATAACAAC ATAGGACACT AAGAAAACAC GGCTTACAGA     300

GAATCCAGCT GACTCTATAA GAGGGGTACT TCTGGAGATT AAAATTATCC GAATCACCTT     360

CCCACTGCGG CTGCTGACGT CAGCGAAAGT CAGAACCGAA AGCGGCGAAG AACCTTCAGA     420

AGAGGAGGAA GCACTTCGAC CTTACAAGAG TTGTTGTCGT TGTTGTTGTC GTTCTCTGGC     480

GGAGAAGCGA GTTTGGATCG CGTTTTCCTC GGAGGCTTCT CGGTCTTCCC CTGTTTCTGC     540

AGCTCAGCCA GGCCCTCGCA AATGGCCTGA AGCTTGGCGT CAACGGCGGA ATGAAGAGGC     600

TAATACTCCC CGAAGTCACC ACCGACGGAG GAACCCTGGT GTCGGAGGTT GGGGAAGTTG     660

AGCCTGGCGA AGTCACCTCG GAGCTTGTAC GCGGCCTTGT GGTACGCCAG AGCGGCTTCC     720

TCGGCGGTGT CGAAGGTTCC CAGCCATAGC CTGGTCCGGA TTCTTCGGGA GTCTAATCTC     780

AGCCACCCAC TTCCCCCCTG AGAAAAGAGA GGAACCACAC TCTCTAAGCC AAAGCAAAAG     840

CAGCAGCAGC AGCA ATG GTT AAA GAC ACA AAG CCT TTA GCC TAT GCT GCC        890
               Met Val Lys Asp Thr Lys Pro Leu Ala Tyr Ala Ala
                 1               5                  10

AAT AAT GGA TAC CAA CAA AAG GGT TCT TCT TTT GAT TTT GAT CCT AGC        938
Asn Asn Gly Tyr Gln Gln Lys Gly Ser Ser Phe Asp Phe Asp Pro Ser
         15                  20                  25

GCT CCT CCA CCG TTT AAG ATT GCA GAA ATC AGA GCT TCA ATA CCA AAA        986
Ala Pro Pro Pro Phe Lys Ile Ala Glu Ile Arg Ala Ser Ile Pro Lys
     30                  35                  40

CAT TGC TGG GTC AAG AAT CCA TGG AGA TCC CTC AGT TAT GTT CTC AGG       1034
His Cys Trp Val Lys Asn Pro Trp Arg Ser Leu Ser Tyr Val Leu Arg
 45                  50                  55                  60

GAT GTG CTT GTA ATT GCT GCA TTG GTG GCT GCA GCA ATT CAC TTC GAC       1082
Asp Val Leu Val Ile Ala Ala Leu Val Ala Ala Ala Ile His Phe Asp
                 65                  70                  75

AAC TGG CTT CTC TGG CTA ATC TAT TGC CCC ATT CAA GGC ACA ATG TTC       1130
Asn Trp Leu Leu Trp Leu Ile Tyr Cys Pro Ile Gln Gly Thr Met Phe
             80                  85                  90
```

```
TGG GCT CTC TTT GTT CTT GGA CAT GAT TGT GGC CAT GGA AGC TTT TCA          1178
Trp Ala Leu Phe Val Leu Gly His Asp Cys Gly His Gly Ser Phe Ser
          95                  100                 105

GAT AGC CCT TTG CTG AAT AGC CTG GTG GGA CAC ATC TTG CAT TCC TCA          1226
Asp Ser Pro Leu Leu Asn Ser Leu Val Gly His Ile Leu His Ser Ser
     110                 115                 120

ATT CTT GTG CCA TAC CAT GGA TGG AGA ATT AGC CAC AGA ACT CAC CAT          1274
Ile Leu Val Pro Tyr His Gly Trp Arg Ile Ser His Arg Thr His His
125                 130                 135                 140

CAA AAC CAT GGA CAC ATT GAG AAG GAT GAG TCA TGG GTT CCA TTA ACA          1322
Gln Asn His Gly His Ile Glu Lys Asp Glu Ser Trp Val Pro Leu Thr
                145                 150                 155

GAG AAG ATT TAC AAG AAT CTA GAC AGC ATG ACA AGA CTC ATT AGA TTC          1370
Glu Lys Ile Tyr Lys Asn Leu Asp Ser Met Thr Arg Leu Ile Arg Phe
            160                 165                 170

ACT GTG CCA TTT CCA TTG TTT GTG TAT CCA ATT TAT TTG TTT TCA AGA          1418
Thr Val Pro Phe Pro Leu Phe Val Tyr Pro Ile Tyr Leu Phe Ser Arg
        175                 180                 185

AGC CCC GGA AAG GAA GGC TCT CAC TTC AAT CCC TAC AGC AAT CTG TTC          1466
Ser Pro Gly Lys Glu Gly Ser His Phe Asn Pro Tyr Ser Asn Leu Phe
    190                 195                 200

CCA CCC AGT GAG AGA AAA GGA ATA GCA ATA TCA ACA CTG TGT TGG GCT          1514
Pro Pro Ser Glu Arg Lys Gly Ile Ala Ile Ser Thr Leu Cys Trp Ala
205                 210                 215                 220

ACC ATG TTT TCT CTG CTT ATC TAT CTC TCA TTC ATA ACT AGT CCA CTT          1562
Thr Met Phe Ser Leu Leu Ile Tyr Leu Ser Phe Ile Thr Ser Pro Leu
                225                 230                 235

CTA GTG CTC AAG CTC TAT GGA ATT CCA TAT TGG ATA TTT GTT ATG TGG          1610
Leu Val Leu Lys Leu Tyr Gly Ile Pro Tyr Trp Ile Phe Val Met Trp
            240                 245                 250

CTG GAC TTT GTC ACA TAC TTG CAT CAC CAT GGT CAC CAC CAG AAA CTG          1658
Leu Asp Phe Val Thr Tyr Leu His His His Gly His His Gln Lys Leu
        255                 260                 265

CCT TGG TAC CGC GGC AAG GAA TGG AGT TAT TTA AGA GGT GGC CTC ACC          1706
Pro Trp Tyr Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr
    270                 275                 280

ACT GTG GAT CGT GAC TAT GGT TGG ATC TAT AAC ATT CAC CAT GAC ATT          1754
Thr Val Asp Arg Asp Tyr Gly Trp Ile Tyr Asn Ile His His Asp Ile
285                 290                 295                 300

GGC ACC CAT GTT ATC CAC CAT CTT TTC CCC CAA ATT CCT CAT TAT CAC          1802
Gly Thr His Val Ile His His Leu Phe Pro Gln Ile Pro His Tyr His
                305                 310                 315
```

```
CTC GTT GAA GCG ACA CAA GCA GCA AAA CCA GTT CTT GGA GAT TAC TAC        1850
Leu Val Glu Ala Thr Gln Ala Ala Lys Pro Val Leu Gly Asp Tyr Tyr
            320                 325                 330

CGT GAG CCA GAA AGA TCT GCG CCA TTA CCA TTT CAT CTA ATA AAG TAT        1898
Arg Glu Pro Glu Arg Ser Ala Pro Leu Pro Phe His Leu Ile Lys Tyr
        335                 340                 345

TTA ATT CAG AGT ATG AGA CAA GAC CAC TTC GTA AGT GAC ACT GGA GAT        1946
Leu Ile Gln Ser Met Arg Gln Asp His Phe Val Ser Asp Thr Gly Asp
    350                 355                 360

GTT GTT TAT TAT CAG ACT GAT TCT CTG CTC CTC CAC TCG CAA CGA GAC        1994
Val Val Tyr Tyr Gln Thr Asp Ser Leu Leu Leu His Ser Gln Arg Asp
365                 370                 375                 380

TGAGTTTCAA ACTTTTTGGG TTATTATTTA TTGGATTCTA GCTACTCAAA TTACTTTTTT      2054

TTTAATGTTA TGTTTTTTGG AGTTTAACGT TTTCTGAACA ACTTGCAAAT TACTTGCATA      2114

GAGAGACATG GAATATTTAT TTGAAATTAG TAAGGTAGTA ATAATAAATT TTGAATTGTC      2174

AGTTTCA                                                                2181
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 380 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
Met Val Lys Asp Thr Lys Pro Leu Ala Tyr Ala Ala Asn Asn Gly Tyr
  1               5                  10                  15

Gln Gln Lys Gly Ser Ser Phe Asp Phe Asp Pro Ser Ala Pro Pro Pro
            20                  25                  30

Phe Lys Ile Ala Glu Ile Arg Ala Ser Ile Pro Lys His Cys Trp Val
        35                  40                  45

Lys Asn Pro Trp Arg Ser Leu Ser Tyr Val Leu Arg Asp Val Leu Val
    50                  55                  60

Ile Ala Ala Leu Val Ala Ala Ala Ile His Phe Asp Asn Trp Leu Leu
65                  70                  75                  80

Trp Leu Ile Tyr Cys Pro Ile Gln Gly Thr Met Phe Trp Ala Leu Phe
                85                  90                  95

Val Leu Gly His Asp Cys Gly His Gly Ser Phe Ser Asp Ser Pro Leu
            100                 105                 110
```

```
Leu Asn Ser Leu Val Gly His Ile Leu His Ser Ser Ile Leu Val Pro
        115                 120                 125

Tyr His Gly Trp Arg Ile Ser His Arg Thr His His Gln Asn His Gly
    130                 135                 140

His Ile Glu Lys Asp Glu Ser Trp Val Pro Leu Thr Glu Lys Ile Tyr
145                 150                 155                 160

Lys Asn Leu Asp Ser Met Thr Arg Leu Ile Arg Phe Thr Val Pro Phe
                165                 170                 175

Pro Leu Phe Val Tyr Pro Ile Tyr Leu Phe Ser Arg Ser Pro Gly Lys
            180                 185                 190

Glu Gly Ser His Phe Asn Pro Tyr Ser Asn Leu Phe Pro Pro Ser Glu
        195                 200                 205

Arg Lys Gly Ile Ala Ile Ser Thr Leu Cys Trp Ala Thr Met Phe Ser
    210                 215                 220

Leu Leu Ile Tyr Leu Ser Phe Ile Thr Ser Pro Leu Leu Val Leu Lys
225                 230                 235                 240

Leu Tyr Gly Ile Pro Tyr Trp Ile Phe Val Met Trp Leu Asp Phe Val
                245                 250                 255

Thr Tyr Leu His His His Gly His His Gln Lys Leu Pro Trp Tyr Arg
            260                 265                 270

Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr Thr Val Asp Arg
        275                 280                 285

Asp Tyr Gly Trp Ile Tyr Asn Ile His His Asp Ile Gly Thr His Val
    290                 295                 300

Ile His His Leu Phe Pro Gln Ile Pro His Tyr His Leu Val Glu Ala
305                 310                 315                 320

Thr Gln Ala Ala Lys Pro Val Leu Gly Asp Tyr Tyr Arg Glu Pro Glu
                325                 330                 335

Arg Ser Ala Pro Leu Pro Phe His Leu Ile Lys Tyr Leu Ile Gln Ser
            340                 345                 350

Met Arg Gln Asp His Phe Val Ser Asp Thr Gly Asp Val Val Tyr Tyr
        355                 360                 365

Gln Thr Asp Ser Leu Leu Leu His Ser Gln Arg Asp
    370                 375                 380
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1675 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Glycine max

(vii) IMMEDIATE SOURCE:

(B) CLONE: pSFD-118bwp

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 169..1530

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
CTGTGGCAAT TTTTCTCTTC TCCTTCTGGT TCTCATCTTT GTGTTCTTCT TTGTTTCTCA      60

CCTTTCTGAG GATTTTTCCA TCTTAGTTCC TGGAGGCACC AGGAACCTGA CCAAATAAAT     120

AAACCTTTTT TTTCTTCTAA TTTTTCTGAA GTTTCATTTT TTAGTCCA ATG GCA ACT      177
                                                      Met Ala Thr
                                                        1

TGG TAT CAT CAG AAA TGT GGC TTG AAG CCT CTT GCT CCA GTA ATT CCT      225
Trp Tyr His Gln Lys Cys Gly Leu Lys Pro Leu Ala Pro Val Ile Pro
  5                  10                  15

AGA CCT AGA ACT GGG GCT GCT TTG TCC AGC ACC TCA AGG GTT GAA TTT      273
Arg Pro Arg Thr Gly Ala Ala Leu Ser Ser Thr Ser Arg Val Glu Phe
 20                  25                  30                  35

TTG GAC ACA AAC AAG GTA GTG GCA GGT CCT AAG TTT CAA CCT TTG AGG      321
Leu Asp Thr Asn Lys Val Val Ala Gly Pro Lys Phe Gln Pro Leu Arg
                 40                  45                  50

TGC AAC CTC AGG GAG AGG AAT TGG GGG CTG AAA GTG AGT GCC CCT TTG      369
Cys Asn Leu Arg Glu Arg Asn Trp Gly Leu Lys Val Ser Ala Pro Leu
             55                  60                  65

AGG GTT GCT TCC ATT GAA GAG GAG CAA AAG AGT GTT GAT TTA ACC AAT      417
Arg Val Ala Ser Ile Glu Glu Glu Gln Lys Ser Val Asp Leu Thr Asn
         70                  75                  80

GGG ACT AAT GGG GTT GAG CAT GAG AAG CTT CCA GAA TTT GAC CCT GGT      465
Gly Thr Asn Gly Val Glu His Glu Lys Leu Pro Glu Phe Asp Pro Gly
 85                  90                  95

GCT CCG CCA CCA TTC AAC TTG GCT GAT ATT AGA GCA GCC ATT CCA AAG      513
Ala Pro Pro Pro Phe Asn Leu Ala Asp Ile Arg Ala Ala Ile Pro Lys
100                 105                 110                 115
```

```
CAT TGC TGG GTG AAG GAC CCT TGG AGG TCC ATG AGC TAT GTG GTG AGG          561
His Cys Trp Val Lys Asp Pro Trp Arg Ser Met Ser Tyr Val Val Arg
                120                 125                 130

GAT GTG ATT GCT GTC TTT GGT TTG GCT GCT GCT GCT GCG TAT CTC AAT          609
Asp Val Ile Ala Val Phe Gly Leu Ala Ala Ala Ala Ala Tyr Leu Asn
            135                 140                 145

AAT TGG TTG GTT TGG CCT CTC TAT TGG GCT GCT CAA GGC ACT ATG TTC          657
Asn Trp Leu Val Trp Pro Leu Tyr Trp Ala Ala Gln Gly Thr Met Phe
        150                 155                 160

TGG GCT CTG TTT GTT CTT GGT CAT GAT TGT GGT CAT GGA AGC TTT TCA          705
Trp Ala Leu Phe Val Leu Gly His Asp Cys Gly His Gly Ser Phe Ser
    165                 170                 175

AAC AAC TCC AAA TTG AAC AGT GTT GTT GGA CAT CTG CTG CAT TCT TCA          753
Asn Asn Ser Lys Leu Asn Ser Val Val Gly His Leu Leu His Ser Ser
180                 185                 190                 195

ATT CTA GTG CCA TAT CAT GGA TGG AGA ATC AGT CAT AGG ACT CAT CAC          801
Ile Leu Val Pro Tyr His Gly Trp Arg Ile Ser His Arg Thr His His
                200                 205                 210

CAA CAT CAT GGT CAT GCT GAA AAT GAT GAA TCA TGG CAT CCG TTG CCT          849
Gln His His Gly His Ala Glu Asn Asp Glu Ser Trp His Pro Leu Pro
            215                 220                 225

GAA AAA TTG TTC AGA AGC TTG GAC ACT GTA ACT CGT ATG TTA AGA TTC          897
Glu Lys Leu Phe Arg Ser Leu Asp Thr Val Thr Arg Met Leu Arg Phe
        230                 235                 240

ACA GCA CCT TTT CCA CTT CTT GCA TTT CCT GTG TAC CTT TTT AGT AGG          945
Thr Ala Pro Phe Pro Leu Leu Ala Phe Pro Val Tyr Leu Phe Ser Arg
    245                 250                 255

AGT CCT GGG AAG ACT GGT TCT CAC TTT GAC CCC AGC AGT GAC TTG TTC          993
Ser Pro Gly Lys Thr Gly Ser His Phe Asp Pro Ser Ser Asp Leu Phe
260                 265                 270                 275

GTT CCC AAT GAA AGA AAA GAT GTT ATT ACT TCC ACA GCT TGT TGG GCT         1041
Val Pro Asn Glu Arg Lys Asp Val Ile Thr Ser Thr Ala Cys Trp Ala
                280                 285                 290

GCT ATG TTG GGA TTG CTT GTT GGA TTG GGG TTT GTA ATG GGT CCA ATT         1089
Ala Met Leu Gly Leu Leu Val Gly Leu Gly Phe Val Met Gly Pro Ile
            295                 300                 305

CAA CTT CTT AAG CTT TAT GGT GTT CCC TAT GTT ATA TTC GTT ATG TGG         1137
Gln Leu Leu Lys Leu Tyr Gly Val Pro Tyr Val Ile Phe Val Met Trp
        310                 315                 320

TTG GAT TTG GTG ACT TAT TTG CAC CAT CAT GGC CAT GAA GAC AAA TTA         1185
Leu Asp Leu Val Thr Tyr Leu His His His Gly His Glu Asp Lys Leu
    325                 330                 335
```

```
CCT TGG TAC CGT GGA AAG GAA TGG AGC TAC CTC AGG GGT GGT CTA ACT          1233
Pro Trp Tyr Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr
340                 345                 350                 355

ACT CTT GAT CGT GAT TAT GGA TGG ATC AAT AAC ATT CAC CAT GAC ATT          1281
Thr Leu Asp Arg Asp Tyr Gly Trp Ile Asn Asn Ile His His Asp Ile
                360                 365                 370

GGC ACT CAT GTC ATT CAT CAC CTA TTT CCT CAA ATT CCA CAC TAT CAC          1329
Gly Thr His Val Ile His His Leu Phe Pro Gln Ile Pro His Tyr His
            375                 380                 385

TTA GTT GAG GCT ACT GAG GCT GCT AAG CCA GTG TTT GGA AAA TAT TAT          1377
Leu Val Glu Ala Thr Glu Ala Ala Lys Pro Val Phe Gly Lys Tyr Tyr
        390                 395                 400

AGA GAA CCA AAG AAA TCA GCA GCA CCT CTT CCT TTT CAC CTT ATT GGG          1425
Arg Glu Pro Lys Lys Ser Ala Ala Pro Leu Pro Phe His Leu Ile Gly
    405                 410                 415

GAA ATA ATA AGG AGC TTC AAG ACT GAC CAT TTT GTT AGT GAC ACG GGG          1473
Glu Ile Ile Arg Ser Phe Lys Thr Asp His Phe Val Ser Asp Thr Gly
420                 425                 430                 435

GAT GTT GTG TAC TAT CAA ACC GAC TCT AAG ATT AAT GGC TCT TCC AAA          1521
Asp Val Val Tyr Tyr Gln Thr Asp Ser Lys Ile Asn Gly Ser Ser Lys
                440                 445                 450

TTA GAG TGAATATTAA AATTCTTTTC TATATAGACA AGAGAGGCTT ATACACAATT          1577
Leu Glu

CTTATTGCTT TAAAGATTGT CTTGAGTTTC TCCGAAAGTT ACTGCACTTA CTTGGAGTTG      1637

AATCCTTCAT TAATAAAGGG ATGGATGGAT CATATAAA                               1675
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 453 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Met Ala Thr Trp Tyr His Gln Lys Cys Gly Leu Lys Pro Leu Ala Pro
  1               5                  10                  15

Val Ile Pro Arg Pro Arg Thr Gly Ala Ala Leu Ser Ser Thr Ser Arg
            20                  25                  30

Val Glu Phe Leu Asp Thr Asn Lys Val Val Ala Gly Pro Lys Phe Gln
        35                  40                  45
```

```
Pro Leu Arg Cys Asn Leu Arg Glu Arg Asn Trp Gly Leu Lys Val Ser
      50                  55                  60

Ala Pro Leu Arg Val Ala Ser Ile Glu Glu Glu Gln Lys Ser Val Asp
 65                  70                  75                  80

Leu Thr Asn Gly Thr Asn Gly Val Glu His Glu Lys Leu Pro Glu Phe
                 85                  90                  95

Asp Pro Gly Ala Pro Pro Pro Phe Asn Leu Ala Asp Ile Arg Ala Ala
            100                 105                 110

Ile Pro Lys His Cys Trp Val Lys Asp Pro Trp Arg Ser Met Ser Tyr
        115                 120                 125

Val Val Arg Asp Val Ile Ala Val Phe Gly Leu Ala Ala Ala Ala Ala
    130                 135                 140

Tyr Leu Asn Asn Trp Leu Val Trp Pro Leu Tyr Trp Ala Ala Gln Gly
145                 150                 155                 160

Thr Met Phe Trp Ala Leu Phe Val Leu Gly His Asp Cys Gly His Gly
                165                 170                 175

Ser Phe Ser Asn Asn Ser Lys Leu Asn Ser Val Val Gly His Leu Leu
            180                 185                 190

His Ser Ser Ile Leu Val Pro Tyr His Gly Trp Arg Ile Ser His Arg
        195                 200                 205

Thr His His Gln His His Gly His Ala Glu Asn Asp Glu Ser Trp His
    210                 215                 220

Pro Leu Pro Glu Lys Leu Phe Arg Ser Leu Asp Thr Val Thr Arg Met
225                 230                 235                 240

Leu Arg Phe Thr Ala Pro Phe Pro Leu Leu Ala Phe Pro Val Tyr Leu
                245                 250                 255

Phe Ser Arg Ser Pro Gly Lys Thr Gly Ser His Phe Asp Pro Ser Ser
            260                 265                 270

Asp Leu Phe Val Pro Asn Glu Arg Lys Asp Val Ile Thr Ser Thr Ala
        275                 280                 285

Cys Trp Ala Ala Met Leu Gly Leu Leu Val Gly Leu Gly Phe Val Met
    290                 295                 300

Gly Pro Ile Gln Leu Leu Lys Leu Tyr Gly Val Pro Tyr Val Ile Phe
305                 310                 315                 320

Val Met Trp Leu Asp Leu Val Thr Tyr Leu His His His Gly His Glu
                325                 330                 335

Asp Lys Leu Pro Trp Tyr Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly
            340                 345                 350
```

```
Gly Leu Thr Thr Leu Asp Arg Asp Tyr Gly Trp Ile Asn Asn Ile His
        355                 360                 365

His Asp Ile Gly Thr His Val Ile His His Leu Phe Pro Gln Ile Pro
    370                 375                 380

His Tyr His Leu Val Glu Ala Thr Glu Ala Ala Lys Pro Val Phe Gly
385                 390                 395                 400

Lys Tyr Tyr Arg Glu Pro Lys Lys Ser Ala Ala Pro Leu Pro Phe His
                405                 410                 415

Leu Ile Gly Glu Ile Ile Arg Ser Phe Lys Thr Asp His Phe Val Ser
            420                 425                 430

Asp Thr Gly Asp Val Val Tyr Tyr Gln Thr Asp Ser Lys Ile Asn Gly
        435                 440                 445

Ser Ser Lys Leu Glu
    450
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 396 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Zea mays

(vii) IMMEDIATE SOURCE:

(B) CLONE: pPCR20

(ix) FEATURE:

(A) NAME/KEY: exon
(B) LOCATION: 31..363

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
GGATCCACGC ATCATCAGAA TCACGGTCAC ATCCACAGGG ACGAGTCATG GCACCCGATC      60

ACGGAGAAGC TGTACCGGCA ACTAGAGCCA CGCACCAAGA AGCTGAGATT CACGGTGCCC     120

TTCCCCCTGC TCGCATTCCC CGTCTACCTC TTGTACAGGA GCCCCGGCAA GCTCGGCTCC     180

CACTTCCTTC CCAGCAGCGA CCTGTTCAGC CCCAAGGAGA AGAGCGACGT CATGGTGTCA     240
```

ACCACCTGCT GGTGCATCAT GCTCGCCTCC CTCCTCGCCA TGGCGTGCGC GTTCGGCCCA 300

CTCCAGGTGC TCAAGATGTA CGGCATCCCA TACCTGGTGT TCGTGATGTG GCTTGACCTG 360

GTGACGTACT TACATCACCA CGGCCACGAT GGATCC 396

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS;

(A) LENGTH: 126 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Zea mays

(vii) IMMEDIATE SOURCE:

(B) CLONE: pPCR20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
His His Gln Asn His Gly His Ile His Arg Asp Glu Ser Trp His Pro
 1               5                  10                  15

Ile Thr Glu Lys Leu Tyr Arg Gln Leu Glu Pro Arg Thr Lys Lys Leu
            20                  25                  30

Arg Phe Thr Val Pro Phe Pro Leu Leu Ala Phe Pro Val Tyr Leu Leu
        35                  40                  45

Tyr Arg Ser Pro Gly Lys Leu Gly Ser His Phe Leu Pro Ser Ser Asp
    50                  55                  60

Leu Phe Ser Pro Lys Glu Lys Ser Asp Val Met Val Ser Thr Thr Cys
65                  70                  75                  80

Trp Cys Ile Met Leu Ala Ser Leu Leu Ala Met Ala Cys Ala Phe Gly
                85                  90                  95

Pro Leu Gln Val Leu Lys Met Tyr Gly Ile Pro Tyr Leu Val Phe Val
            100                 105                 110

Met Trp Leu Asp Leu Val Thr Tyr Leu His His His Gly His
        115                 120                 125
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 472 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: CDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana

(vii) IMMEDIATE SOURCE:

(B) CLONE: pFadx-2 and pYacp7

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
CCTCGAGCTA CGTCAGGGCT AAAACCAGGA ACTGGGCATT GAATGTGGCA ACACCTTTAA      60
CAACTCTTCA GTCTCCATCC GAGGAAGACA GGGAGAGATT CGACCCAGGT GCGCCTCCTC     120
CCTTCAATTT GGCGGATATA AGAGCAGCCA TACCTAAGCA TTGTTGGGTT AAGAATCCAT     180
GGATGTCTAT GAGTTATGTT GTCAGAGATG TTGCTATCGT CTTTGGATTG GCTGCTGTTG     240
CTGCTTACTT CAACAATTGG CTTCTCTGGC CTCTCTACTG GTTCGCTCAA GGAACCATGT     300
TCTGGGCTCT CTTTGTCCTT GGCCATGACT GCGGACATGG TAGCTTCTCG AATGATCCGA     360
GGCTGAACAG TGTGGCTGGT CATCTTCTTC ATTCCTCAAT CCTGGTCCCT TACCATGGCT     420
GGAGGATTAG CCACAGAACT CACCACCAGA ACCATGGTCA TGTCGAGAAT GA              472
```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 156 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana

(vii) IMMEDIATE SOURCE:

(B) CLONE: pFadx-2 and pYacp7

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
Ser Ser Tyr Val Arg Ala Lys Thr Arg Asn Trp Ala Leu Asn Val Ala
1               5                  10                  15

Thr Pro Leu Thr Thr Leu Gln Ser Pro Ser Glu Glu Asp Arg Glu Arg
            20                  25                  30

Phe Asp Pro Gly Ala Pro Pro Pro Phe Asn Leu Ala Asp Ile Arg Ala
        35                  40                  45

Ala Ile Pro Lys His Cys Trp Val Lys Asn Pro Trp Met Ser Met Ser
    50                  55                  60

Tyr Val Val Arg Asp Val Ala Ile Val Phe Gly Leu Ala Ala Val Ala
65                  70                  75                  80

Ala Tyr Phe Asn Asn Trp Leu Leu Trp Pro Leu Tyr Trp Phe Ala Gln
                85                  90                  95

Gly Thr Met Phe Trp Ala Leu Phe Val Leu Gly His Asp Cys Gly His
            100                 105                 110

Gly Ser Phe Ser Asn Asp Pro Arg Leu Asn Ser Val Ala Gly His Leu
        115                 120                 125

Leu His Ser Ser Ile Leu Val Pro Tyr His Gly Trp Arg Ile Ser His
    130                 135                 140

Arg Thr His His Gln Asn His Gly His Val Glu Asn
145                 150                 155
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: misc feature
(B) LOCATION: 1..11
(D) OTHER INFORMATION: /note= "N= INOSINE"

(ix) FEATURE:

(A) NAME/KEY: misc feature
(B) LOCATION: 12..31
(D) OTHER INFORMATION: /note= "N= A OR T OR G OR C"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGGGATCCAC NCAYCAYCAR AAYCAYGGNC A 31

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: misc feature
(B) LOCATION: 1..15
(D) OTHER INFORMATION: /note= "N= INOSINE"

(ix) FEATURE:

(A) NAME/KEY: misc feature
(B) LOCATION: 16..35
(D) OTHER INFORMATION: /note= "N= A OR T OR G OR C"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CGGGATCCRT CRTGNCCRTG RTGRTGNARR TANGT 35

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: misc feature
(B) LOCATION: 1..36
(D) OTHER INFORMATION: /note= "N= INOSINE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TTCGTNNTNG GNCAYGAYTG YGGNCAYGGN CAYGGNAGNT TC 42

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: misc feature
(B) LOCATION: 1..36
(D) OTHER INFORMATION: /note= "N= INOSINE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TTCGTNNTNG GNCAYGAYTG YGGNCAYGGN TCNTTC 36

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION; SEQ ID NO:22:

GGHCAYGAYT GYGGHCAC 18

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GGHCAYGAYT GYGGHCAT 18

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GTACTRTARC CDTGDGTR 18

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GTGCTRTARC CDTGDGTR 18

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GTRCANTARG TRGTRAAYAA YGG 23

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GTRCANTADG TRGTRGADAA YGG 23

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: misc feature
(B) LOCATION: 1..36

(D) OTHER INFORMATION: /note= "N= INOSINE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TTCGTNNTNG GNCAYGAYTG YGGNCAYGGN AGNTTT 36

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: misc feature
(B) LOCATION: 1..36
(D) OTHER INFORMATION: /note= "N= INOSINE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TTCGTNNTNG GNCAYGAYTG YGGNCAYGGN TCNTTT 36

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: misc feature
(B) LOCATION: 1..38
(D) OTHER INFORMATION: /note= "N= INOSINE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GTRCTRTANC CNTGNGTNCA NTANGTAGTG RANAAGGG 38

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: misc feature
(B) LOCATION: 1..38
(D) OTHER INFORMATION: /note= "N= INOSINE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GTRCTRTANC CNTGNGTNCA NTANGTGGTG RANAAGGG 38

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 138 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: misc feature
(B) LOCATION: 1..135
(D) OTHER INFORMATION: /note= "N= INOSINE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GTGGTGNGTN CNGTGNGANA NNCKCCANCC GTGGTANGGN ACNANNANGA ANGANGAGTG 60

NANNANGTGN CCNACNANNG AGTTNANNAN NGGNATNTCN GAGAANGANC CGTGNCCGCA 120

NTCGTGNCCN ANNACGAA 138

## Claims

1. An isolated nucleic acid fragment comprising a nucleic acid sequence encoding a fatty acid desaturase which hybridises to the nucleotide sequence set forth in SEQ ID NO: 1 under one of the following sets of conditions:

   (a) hybridisation in 50 mM Tris-HCl, pH 7.5, 1M NaCl, 1% sodium dodecyl sulfate (SDS), 5% dextran sulfate and 0.1 mg/mL denatured salmon sperm DNA at 50°C and wash twice at room temperature with 2X SSPE 0.1% SDS for 5 minutes and 10 minutes, followed by washing for 5 minutes at 50°C in 0.5X SSPE 0.1% SDS;
   (b) hybridisation in 50 mM Tris-HCl, pH 7.5, 1M NaCl, 1% sodium dodecyl sulfate (SDS), 5% dextran sulfate and 0.1 mg/mL denatured salmon sperm DNA at 50°C and wash twice at room temperature with 2X SSPE, 1% SDS for 5 minutes, then washing for 5 minutes at 50°C in 0.2X SSPE, 1% SDS; or
   (c) hybridisation in 50 mM Tris, pH 7.6, 6X SSC, 5X Denhardt's, 0.5% sodium dodecyl sulfate (SDS), 100 µg denatured calf thymus DNA at 50°C and wash with 6X SSC, 0.5% SDS at room temperature for 15 minutes, repeat with 2X SSC, 0.5% SDS at 45°C for 30 minutes, then repeat twice with 0.2X SSC, 0.5% SDS at 50°C for 30 minutes each.

2. The isolated nucleic acid fragment of claim 1 wherein the nucleic acid identity is 90% or greater to SEQ ID NOS: 1, 4, 6, 8, 10, 12, 14 or 16.

3. An isolated nucleic acid fragment of claim 1 wherein the amino acid identity is 65% or greater to the polypeptide encoded by SEQ ID NO: 14.

4. An isolated nucleic acid fragment of claim 1 comprising a nucleic acid sequence of any one of SEQ ID NOS: 1, 4, 6, 8, 10, 12, 14 or 16.

5. An isolated nucleic acid fragment of claim 1 comprising a nucleic acid sequence encoding a fatty acid desaturase acid sequence with an amino acid sequence encoded by SEQ ID NOS: 1, 4, 6, 8, 10, 12, 14 or 16.

6. An isolated nucleic acid fragment of any one of claims 1 to 5 wherein said fragment is isolated from a plant selected from the group consisting of soybean, oilseed Brassica species, Arabidopsis thaliana and corn.

7. A chimeric gene which comprises heterogenous regulatory and coding sequences not found in nature, comprising a nucleic acid fragment of any of claims 1 to 6, the fragment operably linked to suitable regulatory sequences.

8. Plants containing the chimeric genes of claim 7.

9. A method of producing seed oil containing altered levels of linolenic (18:3) acid comprising:

(a) transforming a plant cell of an oil-producing species with a chimeric gene of claim 7;
(b) growing fertile plants from the transformed plant cells of step (a);
(c) screening progeny seeds from the fertile plants of step (b) for the desired levels of linolenic (18:3) acids; and
(d) processing the progeny seed of step (c) to obtain seed oil containing altered levels of linolenic (18:3) acid.

10. A method of claim 9 wherein said plant cell of an oil-producing species is selected from the group consisting of Arabidopsis thaliana, soybean, oilseed Brassica species, sunflower, cotton, cocoa, peanut, safflower and corn.

11. A method of RFLP mapping with a genomic RFLP marker comprising:

(a) making a cross between two varieties of plants;
(b) making a Southern blot of restriction enzyme digested genomic DNA isolated from several progeny plants resulting from the cross of step (a); and
(c) hybridising the Southern blot with a radiolabelled nucleic acid fragments of claim 1.

12. The isolated genomic DNA of Arabidopsis thaliana identified by accession number ATCC 75167.

13. An isolated cDNA clone which encodes for soybean delta-15 desaturase, comprising the DNA sequence of SEQ ID NO: 10 and identified by accession number ATCC 68874.

14. An isolated cDNA clone which encodes for oilseed Brassica species delta-15 desaturase, comprising the DNA sequence of SEQ ID NO: 6 and identified by accession number ATCC 68854.

**Patentansprüche**

1. Isoliertes Nucleinsäurefragment, umfassend eine Nucleinsäuresequenz, kodierend eine Fettsäure-Desaturase, welche zu der Nucleotidsequenz, angegeben in SEQ ID NO: 1, unter einer der folgenden Gruppen von Bedingungen hybridisiert:

(a) Hybridisierung in 50 mM Tris-HCl, pH 7,5, 1M NaCl, 1% Natriumdodecylsulfat (SDS), 5% Dextransulfat und 0,1 mg/ml DNA von denaturiertem Lachssperma bei 50°C und zweimal Waschen bei Raumtemperatur mit 2X SSPE, 0,1% SDS für 5 Minuten und 10 Minuten, nachfolgend Waschen für 5 Minuten bei 50°C in 0,5X SSPE, 0,1% SDS;

(b) Hybridisierung in 50 mM Tris-HCl, pH 7,5, 1M NaCl, 1% Natriumdodecylsulfat (SDS), 5% Dextransulfat und 0,1 mg/ml DNA von denaturiertem Lachssperma bei 50°C und zweimal Waschen bei Raumtemperatur mit 2X SSPE, 1% SDS für 5 Minuten, dann Waschen für 5 Minuten bei 50°C in 0,2X SSPE, 1% SDS; oder

(c) Hybridisierung in 50 mM Tris, pH 7,6, 6X SSC, 5X Denhardt's, 0,5% Natriumdodecylsulfat (SDS), 100 µg DNA von denaturiertem Kalbsthymus bei 50°C und Waschen mit 6X SSC, 0,5% SDS bei Raumtemperatur für 15 Minuten, Wiederholen mit 2X SSC, 0,5% SDS bei 45°C für 30 Minuten, dann zweimal Wiederholen mit

0,2X SSC, 0,5% SDS bei 50°C für jeweils 30 Minuten.

2. Isoliertes Nucleinsäurefragment nach Anspruch 1, wobei die Nucleinsäureidentität zu 90% oder mehr den SEQ ID NOS: 1, 4, 6, 8, 10, 12, 14 oder 16 entspricht.

3. Isoliertes Nucleinsäurefragment nach Anspruch 1, wobei die Aminosäureidentität zu 65% oder mehr dem Polypeptid, kodiert durch SEQ ID NO: 14, entspricht.

4. Isoliertes Nucleinsäurefragment nach Anspruch 1, umfassend eine Nucleinsäuresequenz mit einer von den SEQ ID NOS: 1, 4, 6, 8, 10, 12, 14 oder 16.

5. Isoliertes Nucleinsäurefragment nach Anspruch 1, umfassend eine Nucleinsäuresequenz, kodierend eine Fettsäure-Desaturase-Säuresequenz mit einer Aminosäuresequenz, kodiert durch SEQ ID NOS: 1, 4, 6, 8, 10, 12, 14 oder 16.

6. Isoliertes Nucleinsäurefragment nach einem der Ansprüche 1 bis 5, wobei das Fragment aus einer Pflanze, ausgewählt aus der Gruppe, bestehend aus Sojabohne, Brassica-Spezies mit Ölsamen, Arabidopsis thaliana und Mais, isoliert wird.

7. Chimäres Gen, welches heterogene regulatorische und kodierende Sequenzen umfaßt, die nicht in der Natur gefunden werden, umfassend ein Nucleinsäurefragment nach einem der Ansprüche 1. bis 6, wobei das Fragment operabel mit geeigneten regulatorischen Sequenzen verknüpft ist.

8. Pflanzen, enthaltend die chimären Gene nach Anspruch 7.

9. Verfahren zur Erzeugung von Samenöl, enthaltend veränderte Gehalte von Linolensäure (18:3), umfassend:

   (a) Tranformieren einer Pflanzenzelle einer ölerzeugenden Spezies mit einem chimären Gen nach Anspruch 7;

   (b) Züchten fruchtbarer Pflanzen aus den transformierten Pflanzenzellen von Schritt (a);

   (c) Screenen der Nachkommensamen von den fruchtbaren Pflanzen von Schritt (b) auf die gewünschten Gehalte von Linolensäuren (18:3); und

   (d) Verarbeiten des Nachkommensamens von Schritt (c), um Samenöl, enthaltend veränderte Gehalte von Linolensäure (18:3), zu erhalten.

10. Verfahren nach Anspruch 9, wobei die Pflanzenzelle einer ölerzeugenden Spezies aus der Gruppe, bestehend aus Arabidopsis thaliana, Sojabohne, Brassica-Spezies mit Ölsamen, Sonnenblume, Baumwolle, Kakao, Erdnuß, Saflor und Mais, ausgewählt ist.

11. Verfahren zum RFLP-Mapping mit einem genomischen RFLP-Marker, umfassend:

   (a) Herstellen einer Kreuzung zwischen zwei Varietäten von Pflanzen;

   (b) Herstellen eines Southern Blot von Restriktionsenzym-digestierter genomischer DNA, isoliert aus verschiedenen Nachkommenpflanzen, entstehend aus der Kreuzung von Schritt (a); und

   (c) Hybridisieren des Southern Blot mit einem radiomarkierten Nucleinsäurefragment nach Anspruch 1.

12. Isolierte genomische DNA von Arabidopsis thaliana, **gekennzeichnet durch** die Zugangsnummer ATCC 75167.

13. Isoliertes cDNA-Klon, welches für Sojabohnen-delta-15-Desaturase kodiert, umfassend die DNA-Sequenz von SEQ ID NO: 10 und **gekennzeichnet durch** die Zugangsnummer ATCC 68874.

14. Isoliertes cDNA-Klon, welches für delta-15-Desaturase von Brassica-Spezies mit Ölsamen kodiert, umfassend die DNA-Sequenz von SEQ ID NO: 6 und **gekennzeichnet durch** die Zugangsnummer ATCC 68854.

**Revendications**

1. Fragment d'acide nucléique isolé, comprenant une séquence d'acide nucléique codant une désaturase d'acide gras qui s'hybride à la séquence de nucléotide indiquée en SEQ ID NO :1 dans l'une des conditions suivantes :

   (a) hybridation dans Tris-HCl 50 mM, pH 7,5, NaCl 1 M, dodécylsulfate de sodium (SDS) à 1%, sulfate de dextran à 5% et 0,1 mg/ml d'ADN de sperme de saumon dénaturé, à 50°C et lavage deux fois à la température ambiante avec 2X SSPE, SDS à 0,1% pendant 5 minutes et 10 minutes, puis lavage pendant 5 minutes à 50°C dans 0,5X SSPE, SDS à 0,1%;
   (b) hybridation dans Tris-HCl 50 mM, pH 7,5, NaCl 1 M, dodécylsulfate de sodium (SDS) à 1%, sulfate de dextran à 5% et 0,1 mg/ml d'ADN de sperme de saumon dénaturé, à 50°C et lavage deux fois à la température ambiante avec 2X SSPE, SDS à 1% pendant 5 minutes, puis lavage pendant 5 minutes à 50°C dans 0,2 x SSPE, SDS à 1%; ou
   (c) hybridation dans Tris 50 mM, pH 7,6, 6X SSC, 5X milieu de Denhardt, dodécylsulfate de sodium (SDS) à 0,5%, et 100µg d'ADN de thymus de veau dénaturé, à 50°C et lavage avec 6X SSC, SDS à 0,5% à la température ambiante pendant 15 minutes, répétition avec 2X SSC, SDS à 0,5% à 45°C pendant 30 minutes, puis répétition deux fois avec 0,2X SSC, SDS à 0,5% à 50°C pendant 30 minutes chaque fois.

2. Fragment d'acide nucléique isolé selon la revendication 1, dans lequel l'identité de l'acide nucléique est à 90% ou plus de SEQ ID NO:1, 4, 6, 8, 10, 12, 14 ou 16.

3. Fragment d'acide nucléique isolé selon la revendication 1, dans lequel l'identité de l'acide nucléique est à 65% ou plus du polypeptide codé par SEQ ID NO:14.

4. Fragment d'acide nucléique isolé selon la revendication 1, comprenant une séquence d'acide nucléique de l'une quelconque parmi SEQ ID NO: 1, 4, 6, 8, 10, 12, 14 ou 16.

5. Fragment d'acide nucléique isolé selon la revendication 1, comprenant une séquence d'acide nucléique codant une séquence de désaturase d'acide gras avec une séquence des acides aminés codée par SEQ ID NO : 1, 4, 6, 8, 10, 12, 14 ou 16.

6. Fragment d'acide nucléique isolé selon l'une quelconque des revendications 1 à 5, dans lequel ledit fragment est isolé d'une plante sélectionnée parmi le groupe consistant en le soya, des espèces de *Brassica* à graines huileuses, *Arabidopsis thaliana* et le maïs.

7. Gène chimérique, qui comprend des séquences de régulation et codantes hétérogènes non trouvées dans la nature, comprenant un fragment d'acide nucléique isolé selon l'une quelconque des revendications 1 à 6, le fragment étant lié de manière fonctionnelle aux séquences de régulation appropriées.

8. Plantes contenant les gènes chimériques selon la revendication 7.

9. Procédé de production d'huile de grain, contenant des taux altérés en acide linolénique (18:3), comprenant :

   (a) la transformation d'une cellule de plante d'une espèce produisant de l'huile avec un gène chimérique selon la revendication 7;
   (b) la croissance en plantes fertiles à partir des cellules de plante transformées de l'étape (a) ;
   (c) le criblage des graines de la descendance des plantes fertiles de l'étape (b) pour le taux souhaité en acide linolénique (18:3), et
   (d) le traitement des graines de la descendance de l'étape (c) pour obtenir une huile de grain contenant des taux altérés en acide linolénique (18:3).

10. Procédé selon la revendication 9, dans lequel ladite cellule de plante d'une espèces produisant de l'huile est sélectionnée parmi le groupe consistant en *Arabidopsis thaliana*, le soya, des espèces de *Brassica* à graines huileuses, le tournesol, le coton, le cacao, l'arachide, le carthame et le maïs.

11. Procédé de préparation d'une carte RFLP avec un marqueur génomique RFLP, comprenant :

    (a) la réalisation d'un croisement entre deux variétés de plantes ;

(b) la préparation d'un transfert de Southern de l'ADN génomique digéré par enzyme de restriction, isolé de plusieurs plantes de la descendance, qui résultent du croisement de l'étape (a), et
(c) l'hybridation du transfert de Southern avec un fragment radiomarqué d'acide nucléique de la revendication 1.

**12.** ADN génomique isolé de *Arabidopsis thaliana*, identifié par le numéro d'accès ATCC 75167.

**13.** Clone d'ADNc isolé, qui code pour la désaturase delta-15 du soya, comprenant la séquence d'ADN de SEQ ID NO:10 et identifié par le numéro d'accès ATCC 68874.

**14.** Clone d'ADNc isolé, qui code pour la désaturase delta-15 d'espèces de *Brassica* à graines huileuses, comprenant la séquence d'ADN de SEQ ID NO :6 et identifié par le numéro d'accès ATCC 68854.

# EXHIBIT G

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 537 178 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication of patent specification: **31.08.94**

(51) Int. Cl.[5]: **C12N 15/53, C11B 1/04, A01H 1/04, C12Q 1/68**

(21) Application number: **91909981.2**

(22) Date of filing: **16.05.91**

(86) International application number:
**PCT/US91/03288**

(87) International publication number:
**WO 91/18985 (12.12.91 91/28)**

# (54) NUCLEOTIDE SEQUENCE OF SOYBEAN STEAROYL-ACP DESATURASE GENE.

(30) Priority: **25.05.90 US 529049**

(43) Date of publication of application:
**21.04.93 Bulletin 93/16**

(45) Publication of the grant of the patent:
**31.08.94 Bulletin 94/35**

(84) Designated Contracting States:
**AT BE CH DE DK ES FR GB GR IT LI LU NL SE**

(56) References cited:
**EP-A- 255 377**
**WO-A-90/12084**
**NL-A- 8 800 794**

**GENE, vol. 72, 1988, Amsterdam NL, pp. 45-50, A. VAN DER KROL**

**BIOTECHNOLOGY, vol. 6, August 1988, New York US, pp. 915-922, M.A.W. HINCHEE et al.**

(73) Proprietor: **E.I. DU PONT DE NEMOURS AND COMPANY**
**1007 Market Street**
**Wilmington Delaware 19898 (US)**

(72) Inventor: **HITZ, William, D.**
**2501 West 18th Street**
**Wilmington, DE 19806 (US)**
Inventor: **YADAV, Narendra, S.**
**127 Jade Drive**
**Wilmington, DE 19810 (US)**

(74) Representative: **Hildyard, Edward Martin et al**
**Frank B. Dehn & Co.**
**Imperial House**
**15-19 Kingsway**
**London WC2B 6UZ (GB)**

**EP 0 537 178 B1**

CHEMICAL ABSTRACTS, vol. 85, no. 17, October 25, 1976, Columbus, Ohio, US; abstract no. 119683, p. 306

JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 257, no. 20, October 25, 1982, BALTIMORE US, T.A. MCKEON & P.K. STUMPF, pp. 12141-12147

TRENDS IN BIOTECHNOLOGY, vol. 5, 1987, Cambridge GB, pp. 40-46, V. KNAUF

BIOTECHNOLOGY, vol. 7, March 1989, New York US, pp. 257-264, S.D. TANKSLEY et al.

PLANT PHYSIOLOGY, vol. 90, 1989, Rockville USA, pp. 760-764, T. CHEESBROUGH

## Description

Soybean oil accounts for about 70% of the 14 billion pounds of edible oil consumed in the United States and is a major edible oil worldwide. It is used in baking, frying, salad dressing, margarine, and a multitude of processed foods. In 1987/88 60 million acres of soybean were planted in the U.S. Soybean is the lowest-cost producer of vegetable oil, which is a by-product of soybean meal. Soybean is agronomically well-adapted to many parts of the U.S. Machinery and facilities for harvesting, storing, and crushing are widely available across the U.S. Soybean products are also a major element of foreign trade since 30 million metric tons of soybeans, 25 million metric tons of soybean meal, and 1 billion pounds of soybean oil were exported in 1987/88. Nevertheless, increased foreign competition has lead to recent declines in soybean acreage and production. The low cost and ready availability of soybean oil provides an excellent opportunity to upgrade this commodity oil into higher value speciality oils to both add value to soybean crop for the U.S. farmer and enhance U.S. trade.

Soybean oil derived from commercial varieties is composed primarily of 11% palmitic (16:0), 4% stearic (18:0), 24% oleic (18:1), 54% linoleic (18:2) and 7% linolenic (18:3) acids. Palmitic and stearic acids are, respectively, 16- and 18-carbon-long saturated fatty acids. Oleic, linoleic and linolenic are 18-carbon-long unsaturated fatty acids containing one, two and three double bonds, respectively. Oleic acid is also referred to as a monounsaturated fatty acid, while linoleic and linolenic acids are also referred to as polyunsaturated fatty acids. The specific performance and health attributes of edible oils is determined largely by their fatty acid composition.

Soybean oil is high in saturated fatty acids when compared to other sources of vegetable oil and contains a low proportion of oleic acid, relative to the total fatty acid content of the soybean seed. These characteristics do not meet important health needs as defined by the American Heart Association.

More recent research efforts have examined the role that monounsaturated fatty acid plays in reducing the risk of coronary heart disease. In the past, it was believed that monounsaturates, in contrast to saturates and polyunsaturates, had no effect on serum cholesterol and coronary heart disease risk. Several recent human clinical studies suggest that diets high in monounsaturated fat may reduce the "bad" (low-density lipoprotein) cholesterol while maintaining the "good" (high-density lipoprotein) cholesterol. [See Mattson et al. (1985) Journal of Lipid Research 26:194-202, Grundy (1986) New England Journal of Medicine 314:745-748, and Mensink et al. (1987) The Lancet 1:122-125, all collectively herein incorporated by reference.] These results corroborate previous epidemiological studies of people living in Mediterranean countries where a relatively high intake of monounsaturated fat and low consumption of saturated fat correspond with low coronary heart disease mortality. [Keys, A., Seven Countries: A Multivariate Analysis of Death and Coronary Heart Disease, Cambridge: Harvard University Press, 1980, herein incorporated by reference.] The significance of monounsaturated fat in the diet was further confirmed by international researchers from seven countries at the Second Colloquim on Monounsaturated Fats held February 26, 1987, in Bethesda, MD, and sponsored by the National Heart, Lung and Blood Institutes [Report, Monounsaturates Use Said to Lower Several Major Risk Factors, Food Chemical News, March 2, 1987, p. 44, herein incorporated by reference.]

Soybean oil is also relatively high in polyunsaturated fatty acids -- at levels in far excess of our essential dietary requirement. These fatty acids oxidize readily to give off-flavors and result in reduced performance associated with unprocessed soybean oil. The stability and flavor of soybean oil is improved by hydrogenation, which chemically reduces the double bonds. However, the need for this processing reduces the economic attractiveness of soybean oil.

A soybean oil low in total saturates and polyunsaturates and high in monounsaturate would provide significant health benefits to the United States population, as well as, economic benefit to oil processors. Soybean varieties which produce seeds containing the improved oil will also produce valuable meal as animal feed.

Another type of differentiated soybean oil is an edible fat for confectionary uses. More than 2 billion pounds of cocoa butter, the most expensive edible oil, are produced worldwide. The U.S. imports several hundred million dollars worth of cocoa butter annually. The high and volatile prices and uncertain supply of cocoa butter have encouraged the development of cocoa butter substitutes. The fatty acid composition of cocoa butter is 26% palmitic, 34% stearic, 35% oleic and 3% linoleic acids. About 72% of cocoa butter's triglycerides have the structure in which saturated fatty acids occupy positions 1 and 3 and oleic acid occupies position 2. Cocoa butter's unique fatty acid composition and distribution on the triglyceride molecule confer on it properties eminently suitable for confectionary end-uses: it is brittle below 27 • C and depending on its crystalline state, melts sharply at 25-30 • C or 35-36 • C. Consequently, it is hard and non-greasy at ordinary temperatures and melts very sharply in the mouth. It is also extremely resistant to

rancidity. For these reasons, producing soybean oil with increased levels of stearic acid, especially in soybean lines containing higher-than-normal levels of palmitic acid, and reduced levels of unsaturated fatty acids is expected to produce a cocoa butter substitute in soybean. This will add value to oil and food processors as well as reduce the foreign import of certain tropical oils.

Only recently have serious efforts been made to improve the quality of soybean oil through plant breeding, especially mutagenesis, and a wide range of fatty acid composition has been discovered in experimental lines of soybean (Table 1). These findings (as well as those with other oilcrops) suggest that the fatty acid composition of soybean oil can be significantly modified without affecting the agronomic performance of a soybean plant. However, there is no soybean mutant line with levels of saturates less than those present in commercial canola, the major competitor to soybean oil as a "healthy" oil.

TABLE 1

| Range of Fatty Acid Percentages Produced by Soybean Mutants | |
|---|---|
| Fatty Acids | Range of Percentages |
| Palmitic Acid | 6-28 |
| Stearic Acid | 3-30 |
| Oleic Acid | 17-50 |
| Linoleic Acid | 35-60 |
| Linolenic Acid | 3-12 |

There are serious limitations to using mutagenesis to alter fatty acid composition. One is unlikely to discover mutations a) that result in a dominant ("gain-of-function") phenotype, b) in genes that are essential for plant growth, and c) in an enzyme that is not rate-limiting and that is encoded by more than one gene. Even when some of the desired mutations are available in soybean mutant lines their introgression into elite lines by traditional breeding techniques will be slow and expensive, since the desired oil compositions in soybean are most likely to involve several recessive genes.

Recent molecular and cellular biology techniques offer the potential for overcoming some of the limitations of the mutagenesis approach, including the need for extensive breeding. Particularly useful technologies are: a) seed-specific expression of foreign genes in transgenic plants [see Goldberg et al. (1989) Cell 56:149-160], b) use of antisense RNA to inhibit plant target genes in a dominant and tissue-specific manner [see van der Krol et al. (1988) Gene 72:45-50], c) transfer of foreign genes into elite commercial varieties of commercial oilcrops, such as soybean [Chee et al. (1989) Plant Physiol. 91:1212-1218; Christou et al. (1989) Proc. Natl. Acad. Sci. U.S.A. 86:7500-7504; Hinchee et al. (1988) Bio/Technology 6:915-922; EPO publication 0 301 749 A2], rapeseed [De Block et al. (1989) Plant Physiol. 91:694-701], and sunflower [Everett et al.(1987) Bio/Technology 5:1201-1204], and d) use of genes as restriction fragment length polymorphism (RFLP) markers in a breeding program, which makes introgression of recessive traits into elite lines rapid and less expensive [Tanksley et al. (1989) Bio/Technology 7:257-264]. However, application of each of these technologies requires identification and isolation of commercially-important genes.

Oil biosynthesis in plants has been fairly well-studied [see Harwood (1989) in Critical Reviews in Plant Sciences, Vol. 8(1) 1-43]. The biosynthesis of palmitic, stearic and oleic acids occur in the plastids by the interplay of three key enzymes of the "ACP track": palmitoyl-ACP elongase, stearoyl-ACP desaturase and acyl-ACP thioesterase. Stearoyl-ACP desaturase introduces the first double bond on stearoyl-ACP to form oleoyl-ACP. It is pivotal in determining the degree of unsaturation in vegetable oils. Because of its key position in fatty acid biosynthesis it is expected to be an important regulatory step. While the enzyme's natural substrate is stearoyl-ACP, it has been shown that it can, like its counterpart in yeast and mammalian cells, desaturate stearoyl-CoA, albeit poorly [McKeon et al. (1982) J. Biol. Chem. 257:12141-12147]. The fatty acids synthesized in the plastid are exported as acyl-CoA to the cytoplasm. At least three different glycerol acylating enzymes (glycerol-3-P acyltransferase, 1-acylglycerol-3-P acyltransferase and diacylglycerol acyltransferase) incorporate the acyl moieties from the cytoplasm into triglycerides during oil biosynthesis. These acyltransferases show a strong, but not absolute, preference for incorporating saturated fatty acids at positions 1 and 3 and monounsaturated fatty acid at position 2 of the triglyceride. Thus, altering the fatty acid composition of the acyl pool will drive by mass action a corresponding change in the fatty acid composition of the oil. Furthermore, there is experimental evidence that, because of this specificity, given the correct composition of fatty acids, plants can produce cocoa butter substitutes [Bafor

et al. (1990) JAOCS 67:217-225].

Based on the above discussion, one approach to altering the levels of stearic and oleic acids in vegetable oils is by altering their levels in the cytoplasmic acyl-CoA pool used for oil biosynthesis. There are two ways of doing this genetically: a) altering the biosynthesis of stearic and oleic acids in the plastid by modulating the levels of stearoyl-ACP desaturase in seeds through either overexpression or antisense inhibition of its gene, and b) converting stearoyl-CoA to oleoyl-CoA in the cytoplasm through the expression of the stearoyl-ACP desaturase in the cytoplasm.

In order to use antisense inhibition of stearoyl-ACP desaturase in the seed, it is essential to isolate the gene(s) or cDNA(s) encoding the target enzyme(s) in the seed, since antisense inhibition requires a high-degree of complementarity between the antisense RNA and the target gene that is expected to be absent in stearoyl-ACP desaturase genes from other species or even in soybean stearoyl-ACP desaturase genes that are not expressed in the seed.

The purification and nucleotide sequences of mammalian microsomal stearoyl-CoA desaturases have been published [Thiede et al. (1986) J. Biol. Chem. 262:13230-13235; Ntambi et al. (1988) J. Biol. Chem. 263:17291-17300; Kaestner et al. (1989) J. Biol. Chem. 264:14755-14761]. However, the plant enzyme differs from them in being soluble, in utilizing a different electron donor, and in its substrate-specificities. The purification and the nucleotide sequences for animal enzymes do not teach how to purify the plant enzyme or isolate a plant gene. The purification of stearoyl-ACP desaturase was reported from safflower seeds [McKeon et al. (1982) J. Biol. Chem. 257:12141-12147]. However, this purification scheme was not useful for soybean, either because the desaturases are different or because of the presence of other proteins such as the soybean seed storage proteins in seed extracts.

The rat liver stearoyl-CoA desaturase protein has been expressed in E. coli [Strittmatter et al. (1988) J. Biol. Chem. 263:2532-2535] but, as mentioned above, its substrate specificity and electron donors are quite distinct from that of the plant.

## SUMMARY OF THE INVENTION

A means to control the levels of saturated and unsaturated fatty acids in edible plant oils has been discovered. Utilizing the soybean seed stearoyl-ACP desaturase cDNA for either the precursor or enzyme, chimeric genes are created and may be utilized to transform various plants to modify the fatty acid composition of the oil produced. Specifically, one aspect of the present invention is a nucleic acid fragment comprising a nucleotide sequence encoding the soybean seed stearoyl-ACP desaturase cDNA corresponding to the nucleotides 1 to 2243 in SEQ ID NO:1, or any nucleic acid fragment substantially homologous therewith. Preferred are those nucleic acid fragments encoding the soybean seed stearoyl-ACP desaturase precursor or the mature soybean seed stearoyl-ACP desaturase enzyme.

Another aspect of this invention involves a chimeric gene capable of transforming a soybean plant cell comprising a nucleic acid fragment encoding the soybean seed stearoyl-ACP desaturase cDNA operably linked to suitable regulatory sequences producing antisense inhibition of soybean seed stearoyl-ACP desaturase in the seed. Preferred are those chimeric genes which incorporate nucleic acid fragments encoding the soybean seed stearoyl-ACP desaturase precursor or the mature soybean seed stearoyl-ACP desaturase enzyme.

Yet another embodiment of the invention involves a method of producing seed oil containing modified levels of saturated and unsaturated fatty acids comprising: (a) transforming a plant cell with a chimeric gene described above, (b) growing sexually mature plants from said transformed plant cells, (c) screening progeny seeds from said sexually mature plants for the desired levels of stearic acid, and (d) crushing said progeny seed to obtain said oil containing modified levels of stearic acid. Preferred plant cells and oils are derived from soybean, rapeseed, sunflower, cotton, cocoa, peanut, safflower, and corn. Preferred methods of transforming such plant cells would include the use of Ti and Ri plasmids of Agrobacterium, electroporation, and high-velocity ballistic bombardment.

## DETAILED DESCRIPTION OF THE INVENTION

The present invention describes a nucleic acid fragment that encodes soybean seed stearoyl-ACP desaturase. This enzyme catalyzes the introduction of a double bond between carbon atoms 9 and 10 of stearoyl-ACP to form oleoyl-ACP. It can also convert stearoyl-CoA into oleoyl-CoA, albeit with reduced efficiency. Transfer of the nucleic acid fragment of the invention, or a part thereof that encodes a functional enzyme, with suitable regulatory sequences into a living cell will result in the production or over-production of stearoyl-ACP desaturase, which in the presence of an appropriate electron donor, such as ferredoxin,

may result in an increased level of unsaturation in cellular lipids, including oil, in tissues when the enzyme is absent or rate-limiting.

Occasionally, reintroduction of a gene or a part thereof into a plant results in the inhibition of both the reintroduced and the endogenous gene, Jorgenson (December, 1990) Trends in Biotechnology 340-344. Therefore, reintroduction of the nucleic acid fragment of the invention is also expected to, in some cases, result in inhibition of the expression of endogenous seed stearoyl-ACP desaturase and would then result in increased level of saturation in seed oil.

Transfer of the nucleic acid fragment of the invention into a soybean plant with suitable regulatory sequences that transcribe the antisense RNA complementary to the mRNA, or its precursor, for seed stearoyl-ACP desaturase may result in the inhibition of the expression of the endogenous stearoyl-ACP desaturase gene and, consequently, in reduced desaturation in the seed oil.

The nucleic acid fragment of the invention can also be used as a restriction fragment length polymorphism marker in soybean genetic studies and breeding programs.

In the context of this disclosure, a number of terms shall be utilized. As used herein, the term "nucleic acid" refers to a large molecule which can be single stranded or double stranded, composed of monomers (nucleotides) containing a sugar, phosphate and either a purine or pyrimidine. A "nucleic acid fragment" is a fraction of a given nucleic acid molecule. In higher plants, deoxyribonucleic acid (DNA) is the genetic material while ribonucleic acid (RNA) is involved in the transfer of the information in DNA into proteins. A "genome" is the entire body of genetic material contained in each cell of an organism. The term "nucleotide sequence" refers to a polymer of DNA or RNA which can be single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases capable of incorporation into DNA or RNA polymers. As used herein, the term "homologous to" refers to the complementarity between the nucleotide sequence of two nucleic acid molecules or between the amino acid sequences of two protein molecules. Estimates of such homology are provided by either DNA-DNA or DNA-RNA hybridization under conditions of stringency as is well understood by those skilled in the art [as described in Hames and Higgins, Eds. (1985) Nucleic Acid Hybridisation, IRL Press, Oxford, U.K.]; or by the comparison of sequence similarity between two nucleic acids or proteins. As used herein, "substantially homologous" refers to nucleic acid molecules which require less stringent conditions of hybridization than those for homologous sequences, and coding DNA sequence which may involve base changes that do not cause a change in the encoded amino acid, or which involve base changes which may alter an amino acid, but not affect the functional properties of the protein encoded by the DNA sequence.

Thus, the nucleic acid fragments described herein include molecules which comprise possible variations of the nucleotide bases derived from deletion, rearrangement, random or controlled mutagenesis of the nucleic acid fragment, and even occasional nucleotide sequencing errors so long as the DNA sequences are substantially homologous.

"Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding) and following (3' non-coding) the coding region. "Stearoyl-ACP desaturase gene" refers to a nucleic acid fragment that expresses a protein with stearoyl-ACP desaturase activity. "Native" gene refers to the gene as found in nature with its own regulatory sequences. "Chimeric" gene refers to a gene that comprises heterogeneous regulatory and coding sequences. "Endogenous" gene refers to the native gene normally found in its natural location in the genome. A "foreign" gene refers to a gene not normally found in the host organism but that is introduced by gene transfer.

"Coding sequence" refers to a DNA sequence that codes for a specific protein and excludes the non-coding sequences. It may constitute an "uninterrupted coding sequence", i.e., lacking an intron, such as in a cDNA or it may include one or more introns bounded by appropriate splice junctions. An "intron" is a sequence of RNA which is transcribed in the primary transcript but which is removed through cleavage and re-ligation of the RNA within the cell to create the mature mRNA that can be translated into a protein.

"Translation initiation codon" and "translation termination codon" refer to a unit of three adjacent nucleotides in a coding sequence that specifies initiation and chain termination, respectively, of protein synthesis (mRNA translation). "Open reading frame" refers to the amino acid sequence encoded between translation initiation and termination codons of a coding sequence.

"RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA" (mRNA) refers to the RNA that is without introns and that can be translated into protein by the cell. "cDNA" refers to a double-stranded DNA that is complementary to and derived from mRNA. "Sense" RNA refers to an RNA transcript that includes the mRNA. "Antisense RNA" refers to an RNA transcript that is complementary to all or part of a

target primary transcript or mRNA and that blocks the expression of a target gene by interfering with the processing, transport and/or translation of its primary transcript or mRNA. The complementarity of an antisense RNA may be with any part of the specific gene transcript, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. In addition, as used herein, antisense RNA may contain regions of ribozyme sequences that may increase the efficacy of antisense RNA to block gene expression. "Ribozyme" refers to a catalytic RNA and includes sequence-specific endoribonucleases.

As used herein, "suitable regulatory sequences" refer to nucleotide sequences located upstream (5'), within, and/or downstream (3') to a coding sequence, which control the transcription and/or expression of the coding sequences, potentially in conjunction with the protein biosynthetic apparatus of the cell. In artificial DNA constructs, regulatory sequences can also control the transcription and stability of antisense RNA.

"Promoter" refers to a DNA sequence in a gene, usually upstream (5') to its coding sequence, which controls the expression of the coding sequence by providing the recognition for RNA polymerase and other factors required for proper transcription. In artificial DNA constructs promoters can also be used to transcribe antisense RNA. Promoters may also contain DNA sequences that are involved in the binding of protein factors which control the effectiveness of transcription initiation in response to physiological or developmental conditions. It may also contain enhancer elements. An "enhancer" is a DNA sequence which can stimulate promoter activity. It may be an innate element of the promoter or a heterologous element inserted to enhance the level and/or tissue-specificity of a promoter. "Constitutive promoters" refers to those that direct gene expression in all tissues and at all times. "Tissue-specific" or "development-specific" promoters as referred to herein are those that direct gene expression almost exclusively in specific tissues, such as leaves or seeds, or at specific development stages in a tissue, such as in early or late embryogenesis, respectively. "Inducible promoters" refers to those that direct gene expression in response to an external stimulus, such as light, heat-shock and chemical.

The term "expression", as used herein, is intended to mean the production of a functional end-product. In the case of expression or overexpression of the stearoyl-ACP desaturase genes it involves transcription of the gene and translation of the mRNA into precursor or mature stearoyl-ACP desaturase proteins. In the case of antisense inhibition it refers to the production of antisense RNA transcripts capable of preventing the expression of the target protein. "Overexpression" refers to the production of a gene product in transgenic organisms that exceeds levels of production in normal or non-transformed organisms.

The "3' non-coding sequences" refers to that the DNA sequence portion of a gene that contains a polyadenylation signal and any other regulatory signal capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor.

"Mature" protein refers to a functional desaturase enzyme without its transit peptide. "Precursor" protein refers to the mature protein with a native or foreign transit peptide. "Transit" peptide refers to the amino terminal extension of a polypeptide, which is translated in conjunction with the polypeptide forming a precursor peptide and which is required for its uptake by plastids of a cell.

"Transformation" herein refers to the transfer of a foreign gene into the genome of a host organism and its genetically stable inheritance. "Restriction fragment length polymorphism" refers to different sized restriction fragment lengths due to altered nucleotide sequences in or around variant forms of genes, and may be abbreviated as "RFLP". "Fertile" refers to plants that are able to propagate sexually.

Purification of Soybean Seed Stearoyl-ACP Desaturase

Stearoyl-ACP desaturase protein was purified to near-homogeneity from the soluble fraction of extracts made from developing soybean seeds following its chromatography on Blue Sepharose, anion-exchange, alkyl-ACP sepharose, and chromatofocussing on Mono P (Pharmacia). Because of the lability of the enzyme during purification, the nearly homogenous preparation is purified only ca. a few hundred-fold; the basis of this lability is not understood. Chromatofocussing resolved the enzyme into two peaks of activity: the peak that eluted earlier, with an apparent pI of ca. 6, had a higher specific-activity than the peak eluting later, with an apparent pI of ca. 5.7. The native molecular weight of the purified enzyme was estimated by gel filtration to be ca. 65 kD. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of the purified desaturase preparation showed it to be a polypeptide of ca. 38 kD, which suggests that the native enzyme is a dimer. A smaller polypeptide is occasionally observed in varying amounts resulting in a doublet in some preparations. This appears to be due to a proteolytic breakdown of the larger one, since the level of the smaller one increases during storage. However, it cannot be ruled out that the enzyme could also be a heterodimer or that there are different-sized isozymes.

A highly purified desaturase preparation was resolved on SDS-PAGE, electrophoretically transfered onto Immobilon®-P membrane (Millipore), and stained with Coomassie blue. The ca. 38 kD protein on the Immobilon®-P was cut out and used to make polyclonal antibody in mice.

A $C_4$ reverse-phase HPLC column was used to further purify the enzyme that eluted earlier in chromatofocussing. The major protein peak was homogeneous for the ca. 38 kD polypeptide. It was used for determining the N-terminal sequence: Arg-Ser-Gly-Ser-Lys-Glu-Val-Glu-Asn-Ile-Lys-Lys-Pro-Phe-Thr-Pro (SEQ ID NO:3).

Cloning of Soybean Seed Stearoyl-ACP Desaturase cDNA

Based on the N-terminal sequence of the purified desaturase protein, a set of eight degenerate 35 nucleotide-long oligonucleotides was designed for use as a hybridization probe. The design took into account the codon usage in selected soybean seed genes and used five deoxyinosines at selected positions of ambiguity. The probe, following radiolabeling, was used to screen a cDNA expression library made in Lambda ZAP vector from poly $A^+$ RNA from 20-day old developing soybean seeds. Six positively-hybridizing plaques were subjected to plaque purification. Sequences of the pBluescript (Stratagene) vector, including the cDNA inserts, from each of six purified phages were excised in the presence of a helper phage and the resultant phagemids used to infect E. coli cells resulting in a double-stranded plasmids, pDS1 to pDS6.

The cDNA insert in plasmid pDS1 is flanked at one end (the 5' end of the coding sequence) by the unique Eco RI site and at its other end by the unique Hind III site. Both Eco RI and the Hind III sites are from the vector, pBluescript. The nucleotide sequence of the cDNA insert in pDS1 revealed an open reading frame for 402 amino acids that included the mature protein's N-terminal sequence 43 amino acid residues from the N-terminus of the open reading frame (SEQ ID NO:1). At least part of this "presequence" is the transit peptide required for precursor import into the chloroplast. Although there are four methionines in this presequence that are in-frame with the mature protein sequence, the most likely N-terminal residue is methionine at position -32 (with the N-terminal Arg of mature protein being referred to as +1) since: a) the N-terminal methionine in the transit peptide sequences for all known chloroplast precursor proteins, with only one exception, is followed by alanine, and b) the methionine at position -5 is too close to the N-terminus of the mature protein to be the initiating codon for the transit peptide (the smallest transit sequence found thus far is 31 amino acids long). Thus, it can be deduced that the desaturase precursor protein consists of a 32-amino acid long transit peptide and a 359-amino acid long mature protein. Based on fusion-protein studies in which the C-terminus of foreign proteins is fused either to the desaturase precursor at position -10 (Ser) or to the mature desaturase protein at position +10 (Ile), the N-terminus of a functional stearoyl-ACP desaturase enzyme can range at least ± 10 amino acids from Arg at position +1 (SEQ ID NO:1).

The restriction maps of all six plasmids, though not identical, showed a common 0.7 kb Bgl II fragment found within the coding region of the precursor for stearoyl-ACP desaturase in pDS1. This strongly suggests that all six clones encode the stearoyl-ACP desaturase. The partial restriction maps of plasmids pDS1, pDS5 and pDS6 appear to be the identical. The inserts in pDS2 and pDS3, which differ in their physical maps from each other as well as from that of pDS1, were partially sequenced. Their partial nucleotide sequences, including 262 nucleotides from the 3' non-coding region, were identical to that in pDS1.

Of the several cDNA clones isolated from the soybean cDNA library using pDS1 as hybridization probe, five were sequenced in the 3' non-coding sequence and their sequences compared to that of SEQ ID NO:1. The results are summarized below:

| Clone # | Sequence correspondence to SEQ ID NO:1 | Percent Identity |
|---|---|---|
| 1 | 1291-1552 | 100 |
| 2 | 1291-1394 | 100 |
| 3 | 1285-1552 | 100 |
| 4 | 1285-1552 | 100 |
| 5 | 1298-1505 | 91 |

Thus, while the claimed sequence (SEQ ID NO:1) most likely represents the predominantly-expressed stearoyl-ACP desaturase gene in soybean seed, there is at least one other stearoyl-ACP desaturase gene that is 91% homologous at the nucleotide level to the claimed sequence. The partial sequence of clone #5

is shown in SEQ ID NO:2.

As expected, comparison of the deduced amino-acid sequences for soybean stearoyl-ACP desaturase and the rat microsomal stearoyl-CoA desaturases did not reveal any significant homology.

In vitro recombinant DNA techniques were used to make two fusion proteins:

a) a recombinant plasmid pGEXB that encodes a ca. 66 kD fusion protein consisting of a 28 kD glutathione-S-transferase (GST) protein fused at its C-terminus to the ca. 38 kD desaturase precursor protein at amino acid residue -10 from the N-terminus of the mature enzyme (Arg, +1) (SEQ ID NO:1). Extracts of E. coli cells harboring pGEXB, grown under conditions that induce the synthesis of the fusion protein, show stearoyl-ACP desaturase activity and expression of a ca. 66 kD fusion protein that cross-reacts with antibody made against soybean stearoyl-ACP desaturase and that binds to glutathione-agarose affinity column. The affinity column can be used to purify the fusion protein to near-homogeneity in a single step. The desaturase moiety can be cleaved off in the presence of thrombin and separated from the GST by re-chromatography on the glutathione-agarose column; and

b) a recombinant plasmid, pNS2, that encodes a ca. 42 kD fusion protein consisting of 4 kD of the N-terminus of β-galactosidase fused at its C-terminus to the amino acid residue at position +10 (Ile) from the N-terminus of the mature desaturase protein (Arg, +1) (SEQ ID NO:1). Extract of E. coli cells harboring pNS2 express a ca. 42 kD protein that cross-reacts with antibody made against soybean stearoyl-ACP desaturase and show stearoyl-ACP desaturase activity.

E. coli (pGEXB) can be used to purify the stearoyl-ACP desaturase for use in structure-function studies on the enzyme, in immobilized cells or in extracellular desaturations [see Ratledge et al. (1984) Eds., Biotechnology for the Oils and Fats Industry, American Oil Chemists' Society]. E. coli (pNS2) can be used to express the desaturase enzyme in vivo. However, for in vivo function it may be necessary to introduce an electron donor, such as ferredoxin and NADPH:ferredoxin reductase. The ferredoxin gene has been cloned from a higher plant [Smeekens et al. (1985) Nucleic Acids Res. 13:3179-3194] and human ferredoxin has been expressed in E. coli [Coghlan et al. (1989) Proc. Natl. Acad. Sci. USA, 86:835-839]. Alternatively, one skilled in the art can express the mature protein in microorganisms using other expression vectors described in the art [Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press; Milman (1987) Meth. Enzymol. 153:482-491; Duffaud et al. (1987) Meth. Enzymol. 153:492-507; Weinstock (1987) Meth. Enzymol. 154:156-163; E.P.O. Publication 0 295 959 A2).

The fragment of the instant invention may be used, if desired, to isolate substantially homologous stearoyl-ACP desaturase cDNAs and genes, including those from plant species other than soybean. Isolation of homologous genes is well-known in the art. Southern blot analysis reveals that the soybean cDNA for the enzyme hybridizes to several, different-sized DNA fragments in the genomic DNA of tomato, rapeseed (Brassica napus), soybean, corn (a monocotyledenous plant) and Arabidopsis (which has a very simple genome). The Southern blot of corn DNA reveals that the soybean cDNA can also hybridize non-specifically, which may make the isolation of the corn gene more difficult. Although we do not know how many different genes or "pseudogenes" (non-functional genes) are present in any plant, it is expected to be more than one, since stearoyl-ACP desaturase is an important enzyme. Moreover, plants that are amphidiploid (that is, derived from two progenitor species), such as soybean, rapeseed (B. napus), and tobacco will have genes from both progenitor species.

The nucleic acid fragment of the instant invention encoding soybean seed stearoyl-ACP desaturase cDNA, or a coding sequence derived from other cDNAs or genes for the enzyme, with suitable regulatory sequences, can be used to overexpress the enzyme in transgenic soybean as well as other transgenic species. Such a recombinant DNA construct may include either the native stearoyl-ACP desaturase gene or a chimeric gene. One skilled in the art can isolate the coding sequences from the fragment of the invention by using and/or creating sites for restriction endonucleases, as described in Sambrook et al. [(1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press]. Of particular utility are sites for Nco I (5'-CCATGG-3') and Sph I (5'-GCATGC-3') that allow precise removal of coding sequences starting with the initiating codon ATG. The fragment of invention has a Nco I recognition sequence at nucleotide positions 1601-1606 (SEQ ID NO:1) that is 357 bp after the termination codon for the coding sequence. For isolating the coding sequence of stearoyl-ACP desaturase precursor from the fragment of the invention, an Nco I site can be engineered by substituting nucleotide A at position 69 with C. This will allow isolation of the 1533 bp Nco I fragment containing the precursor coding sequence. The expression of the mature enzyme in the cytoplasm is expected to desaturate stearoyl-CoA to oleoyl-CoA. For this it may be necessary to also express the mature ferredoxin in the cytoplasm, the gene for which has been cloned from plants [Smeekens et al. (1985) Nucleic Acids Res. 13:3179-3194]. For isolating the coding sequence for the mature protein, a restriction site can be engineered near nucleotide position 164. For example, substituting nucleotide G with nucleotide C at position 149 or position 154 would result in the

creation of Nco I site or Sph I site, respectively. This will allow isolation of a 1453 bp Nco I fragment or a 1448 bp Sph I-Nco I fragment, each containing the mature protein sequence. Based on fusion protein studies, the N-terminus of the mature stearoyl-ACP desaturase enzyme is not critical for enzyme activity.

Antisense RNA has been used to inhibit plant target genes in a dominant and tissue-specific manner [see van der Krol et al. (1988) Gene 72:45-50; Ecker et al. (1986) Proc. Natl. Acad. Sci. USA 83:5372-5376; van der Krol et al. (1988) Nature 336:866-869; Smith et al. (1988) Nature 334:724-726; Sheehy et al. (1988) Proc. Natl. Acad. Sci. USA 85:8805-8809; Rothstein et al. (1987) Proc. Natl. Acad. Sci. USA 84:8439-8443; Cornelissen et al. (1988) Nucl. Acids Res. 17:833-843; Cornelissen (1989) Nucl. Acid Res. 17:7203-7209; Robert et al. (1989) Plant Mol. Biol. 13:399-409].

The use of antisense inhibition of the seed enzyme would require isolation of the coding sequence for genes that are expressed in the target tissue of the target plant. Thus, it will be more useful to use the fragment of the invention to screen seed-specific cDNA libraries, rather than genomic libraries or cDNA libraries from other tissues, from the appropriate plant for such sequences. Moreover, since there may be more than one gene encoding seed stearoyl-ACP desaturase, it may be useful to isolate the coding sequences from the other genes from the appropriate crop. The genes that are most highly expressed are the best targets for antisense inhibition. The level of transcription of different genes can be studied by known techniques, such as run-off transcription.

For expressing antisense RNA in soybean seed from the fragment of the invention, the entire fragment of the invention (that is, the entire cDNA for soybean stearoyl-ACP desaturase from the unique Eco RI to Hind III sites in plasmid pDS1) may be used. There is evidence that the 3' non-coding sequences can play an important role in antisense inhibition [Ch'ng et al. (1989) Proc. Natl. Acad. Sci. USA 86:10006-10010]. There have also been examples of using the entire cDNA sequence for antisense inhibition [Sheehy et al. (1988) Proc. Natl. Acad. Sci. USA 89:8439-8443]. The Hind III and Eco RI sites can be modified to facilitate insertion of the sequences into suitable regulatory sequences in order to express the antisense RNA.

A preferred host soybean plant for the antisense RNA inhibition of stearoyl-ACP desaturase for producing a cocoa butter substitute in soybean seed oil is a soybean plant containing higher-than-normal levels of palmitic acid, such as A19 double mutant, which is being commercialized by Iowa State University Research Foundation, Inc. (315 Beardshear, Ames, Iowa 50011).

A preferred class of heterologous hosts for the expression of the coding sequence of stearoyl-ACP desaturase precursor or the antisense RNA are eukaryotic hosts, particularly the cells of higher plants. Particularly preferred among the higher plants are the oilcrops, such as soybean (Glycine max), rapeseed (Brassica napus, B. campestris), sunflower (Helianthus annus), cotton (Gossypium hirsutum), corn (Zea mays), cocoa (Theobroma cacao), and peanut (Arachis hypogaea). Expression in plants will use regulatory sequences functional in such plants.

The expression of foreign genes in plants is well-established [De Blaere et al. (1987) Meth. Enzymol. 153:277-291]. The origin of promoter chosen to drive the expression of the coding sequence or the antisense RNA is not critical as long as it has sufficient transcriptional activity to accomplish the invention by increasing or decreasing, respectively, the level of translatable mRNA for stearoyl-ACP desaturase in the desired host tissue. Preferred promoters include strong plant promoters (such as the constitutive promoters derived from Cauliflower Mosaic Virus that direct the expression of the 19S and 35S viral transcripts [Odell et al. (1985) Nature 313:810-812; Hull et al. (1987) Virology 86:482-493]), small subunit of ribulose 1,5-bisphosphate carboxylase [Morelli et al. (1985) Nature 315:200; Broglie et al. (1984) Science 224:838; Hererra-Estrella et al. (1984) Nature 310:115; Coruzzi et al. (1984) EMBO J. 3:1671; Faciotti et al. (1985) Bio/Technology 3:241], maize zein protein [Matzke et al. (1984) EMBO J. 3:1525], and chlorophyll a/b binding protein [Lampa et al. (1986) Nature 316:750-752].

Depending upon the application, it may be desirable to select inducible promoters and/or tissue- or development-specific promoters. Such examples include the light-inducible promoters of the small subunit of ribulose 1,5-bisphosphate carboxylase genes (if the expression is desired in tissues with photosynthetic function).

Particularly preferred tissue-specific promoters are those that allow seed-specific expression. This may be especially useful, since seeds are the primary source of vegetable oils and also since seed-specific expression will avoid any potential deleterious effect in non-seed tissues. Examples of seed-specific promoters include but are not limited to the promoters of seed storage proteins, which can represent up to 90% of total seed protein in many plants. The seed storage proteins are strictly regulated, being expressed almost exclusively in seeds in a highly tissue-specific and stage-specific manner [Higgins et al. (1984) Ann. Rev. Plant Physiol. 35:191-221; Goldberg et al. (1989) Cell 56:149-160]. Moreover, different seed storage proteins may be expressed at different stages of seed development.

Expression of seed-specific genes has been studied in great detail [see reviews by Goldberg et al. (1989) Cell 56:149-160 and Higgins et al. (1984) Ann. Rev. Plant Physiol. 35:191-221]. There are currently numerous examples for seed-specific expression of seed storage protein genes in transgenic dicotyledonous plants. These include genes from dicotyledonous plants for bean β-phaseolin [Sengupta-Gopalan et al. (1985) Proc. Natl. Acad. Sci. USA 82:3320-3324; Hoffman et al. (1988) Plant Mol. Biol. 11:717-729], bean lectin [Voelker et al. (1987) EMBO J. 6: 3571-3577], soybean lectin [Okamuro et al. (1986) Proc. Natl. Acad. Sci. USA 83: 8240-8244], soybean kunitz trypsin inhibitor [Perez-Grau et al. (1989) Plant Cell 1:095-1109], soybean β-conglycinin [Beachy et al. (1985) EMBO J. 4:3047-3053; Barker et al. (1988) Proc. Natl. Acad. Sci. USA 85:458-462; Chen et al. (1988) EMBO J. 7:297-302; Chen et al. (1989) Dev. Genet. 10:112-122; Naito et al. (1988) Plant Mol. Biol. 11:109-123], pea vicilin [Higgins et al. (1988) Plant Mol. Biol. 11:683-695], pea convicilin [Newbigin et al. (1990) Planta 180:461], pea legumin [Shirsat et al. (1989) Mol. Gen. Genetics 215:326]; rapeseed napin [Radke et al. (1988) Theor. Appl. Genet. 75:685-694] as well as genes from monocotyledonous plants such as for maize 15-kD zein [Hoffman et al. (1987) EMBO J. 6:3213-3221], and barley β-hordein [Marris et al. (1988) Plant Mol. Biol. 10:359-366] and wheat glutenin [Colot et al. (1987) EMBO J. 6:3559-3564]. Moreover, promoters of seed-specific genes operably linked to heterologous coding sequences in chimeric gene constructs also maintain their temporal and spatial expression pattern in transgenic plants. Such examples include Arabidopsis thaliana 2S seed storage protein gene promoter to express enkephalin peptides in Arabidopsis and B. napus seeds [Vandekerckhove et al. (1989) Bio/Technology 7:929-932], bean lectin and bean β-phaseolin promoters to express luciferase [Riggs et al. (1989) Plant Sci. 63:47-57], and wheat glutenin promoters to express chloramphenicol acetyl transferase [Colot et al. (1987) EMBO J. 6:3559-3564].

Of particular use in the expression of the nucleic acid fragment of the invention will be the heterologous promoters from several extensively-characterized soybean seed storage protein genes such as those for the Kunitz trypsin inhibitor [Jofuku et al. (1989) Plant Cell 1:1079-1093; Perez-Grain et al. (1989) Plant Cell 1:1095-1109], glycinin [Nielson et al. (1989) Plant Cell 1:313-328], β-conglycinin [Harada et al. (1989) Plant Cell 1:415-425]. Promoters of genes for α- and β-subunits of soybean β-conglycinin storage protein will be particularly useful in expressing the mRNA or the antisense RNA to stearoyl-ACP desaturase in the cotyledons at mid- to late-stages of seed development [Beachy et al. (1985) EMBO J. 4:3047-3053; Barker et al. (1988) Proc. Natl. Acad. Sci. USA 85:458-462; Chen et al. (1988) EMBO J. 7:297-302; Chen et al. (1989) Dev. Genet. 10:112-122; Naito et al. (1988) Plant Mol. Biol. 11:109-123] in transgenic plants, since: a) there is very little position effect on their expression in transgenic seeds, and b) the two promoters show different temporal regulation: the promoter for the α-subunit gene is expressed a few days before that for the β-subunit gene; this is important for transforming rapeseed where oil biosynthesis begins about a week before seed storage protein synthesis [Murphy et al. (1989) J. Plant Physiol. 135:63-69].

Also of particular use will be promoters of genes expressed during early embryogenesis and oil biosynthesis The native regulatory sequences, including the native promoter, of the stearoyl-ACP desaturase gene expressing the nucleic acid fragment of the invention can be used following its isolation by those skilled in the art. Heterologous promoters from other genes involved in seed oil biosynthesis, such as those for B. napus isocitrate lyase and malate synthase [Comai et al. (1989) Plant Cell 1:293-300], Arabidopsis ACP [Post-Beittenmiller et al. (1989) Nucl. Acids Res. 17:1777], B. napus ACP [Safford et al. (1988) Eur. J. Biochem. 174:287-295], B. campestris ACP [Rose et al. (1987) Nucl. Acids Res. 15:7197] may also be used. The partial protein sequences for the relatively-abundant enoyl-ACP reductase and acetyl-CoA carboxylase are published [Slabas et al. (1987) Biochim. Biophys. Acta 877:271-280; Cottingham et al. (1988) Biochim. Biophys. Acta 954: 201-207] and one skilled in the art can use these sequences to isolate the corresponding seed genes with their promoters.

Proper level of expression of stearoyl-ACP mRNA or antisense RNA may require the use of different chimeric genes utilizing different promoters. Such chimeric genes can be transfered into host plants either together in a single expression vector or sequentially using more than one vector.

It is envisioned that the introduction of enhancers or enhancer-like elements into either the native stearoyl-ACP desaturase promoter or into other promoter constructs will also provide increased levels of primary transcription for antisense RNA or in RNA for stearoyl-ACP desaturase to accomplish the inventions. This would include viral enhancers such as that found in the 35S promoter [Odell et al. (1988) Plant Mol. Biol. 10:263-272], enhancers from the opine genes (Fromm et al. (1989) Plant Cell 1:977-984), or enhancers from any other source that result in increased transcription when placed into a promoter operably linked to the nucleic acid fragment of the invention.

Of particular importance is the DNA sequence element isolated from the gene for the α-subunit of β-conglycinin that can confer 40-fold seed-specific enhancement to a constitutive promoter [Chen et al. (1988) EMBO J. 7:297-302; Chen et al. (1989) Dev. Genet. 10:112-122]. One skilled in the art can readily isolate

this element and insert it within the promoter region of any gene in order to obtain seed-specific enhanced expression with the promoter in transgenic plants. Insertion of such an element in any seed-specific gene that is expressed at different times than the β-conglycinin gene will result in expression in transgenic plants for a longer period during seed development.

The invention can also be accomplished by a variety of other methods to obtain the desired end. In one form, the invention is based on modifying plants to produce increased levels of stearoyl-ACP desaturase by virtue of having significantly larger numbers of copies of either the wild-type or a stearoyl-ACP desaturase gene from a different soybean tissue in the plants. This may result in sufficient increases in stearoyl-ACP desaturase levels to accomplish the invention.

Any 3' non-coding region capable of providing a polyadenylation signal and other regulatory sequences that may be required for the proper expression of the stearoyl-ACP desaturase coding region can be used to accomplish the invention. This would include the native 3' end of the substantially homologous soybean stearoyl-ACP desaturase gene(s), the 3' end from a heterologous stearoyl-ACP desaturase gene, the 3' end from viral genes such as the 3' end of the 35S or the 19S cauliflower mosaic virus transcripts, the 3' end from the opine synthesis genes, the 3' ends of ribulose 1,5-bisphosphate carboxylase or chlorophyll a/b binding protein, or 3' end sequences from any source such that the sequence employed provides the necessary regulatory information within its nucleic acid sequence to result in the proper expression of the promoter/stearoyl-ACP desaturase coding region combination to which it is operably linked. There are numerous examples in the art that teach the usefulness of different 3' non-coding regions.

Various methods of transforming cells of higher plants according to the present invention are available to those skilled in the art (see EPO publications 0 295 959 A2 and 0 318 341 A1). Such methods include those based on transformation vectors based on the Ti and Ri plasmids of Agrobacterium spp. It is particularly preferred to use the binary type of these vectors. Ti-derived vectors transform a wide variety of higher plants, including monocotyledonous and dicotyledonous plants, such as soybean, cotton and rape [Pacciotti et al.(1985) Bio/Technology 3:241; Byrne et al. (1987) Plant Cell, Tissue and Organ Culture 8:3; Sukhapinda et al. (1987) Plant Mol. Biol. 8:209-216; Lorz et al. (1985) Mol. Gen. Genet. 199:178; Potrykus (1985) Mol. Gen. Genet. 199:183]. Other transformation methods are available to those skilled in the art, such as direct uptake of foreign DNA constructs [see EPO publication 0 295 959 A2], techniques of electroporation [see Fromm et al. (1986) Nature (London) 319:791] or high-velocity ballistic bombardment with metal particles coated with the nucleic acid constructs [see Kline et al. (1987) Nature (London) 327:70]. Once transformed the cells can be regenerated by those skilled in the art.

Of particular relevance are the recently described methods to transform foreign genes into commercially important crops, such as rapeseed [see De Block et al. (1989) Plant Physiol. 91:694-701], sunflower [Everett et al. (1987) Bio/Technology 5:1201], and soybean [McCabe et al. (1988) Bio/Technology 6:923; Hinchee et al. (1988) Bio/Technology 6:915; Chee et al. (1989) Plant Physiol. 91:1212-1218; Christou et al. (1989) Proc. Natl. Acad. Sci USA 86:7500-7504; EPO Publication 0 301 749 A2].

The use of restriction fragment length polymorphism (RFLP) markers in plant breeding has been well-documented in the art [see Tanksley et al. (1989) Bio/Technology 7:257-264]. The nucleic acid fragment of the invention has been mapped to four different loci on a soybean RFLP map [Tingey et al. (1990) J. Cell Biochem., Supplement 14E p. 291, abstract R153]. It can thus be used as a RFLP marker for traits linked to these mapped loci. More preferably these traits will include altered levels of stearic acid. The nucleic acid fragment of the invention can also be used to isolate the stearoyl-ACP desaturase gene from variant (including mutant) soybeans with altered stearic acid levels. Sequencing of these genes will reveal nucleotide differences from the normal gene that cause the variation. Short oligonucleotides designed around these differences may be used as hybridization probes to follow the variation in stearic and oleic acids. Oligonucleotides based on differences that are linked to the variation may be used as molecular markers in breeding these variant oil traits.

SEQ ID NO:1 represents the nucleotide sequence of a soybean seed stearoyl-ACP desaturase cDNA and the translation reading frame that includes the open reading frame for the soybean seed stearoyl-ACP desaturase. The nucleotide sequence reads from 5' to 3'. Three letter codes for amino acids are used as defined by the Commissioner, 1114 OG 29 (May 15, 1990) incorporated by reference herein. Nucleotide 1 is the first nucleotide of the cDNA insert after the EcoRI cloning site of the vector and nucleotide 2243 is the last nucleotide of the cDNA insert of plasmid pDS1 which encods the soybean seed stearoyl-ACP desaturase. Nucleotides 70 to 72 are the putative translation initiation codon, nucleotides 166 to 168 are the codon for the N-terminal amino acid of the purified enzyme, nucleotides 1243 to 1245 are the termination codon, nucleotides 1 to 69 are the 5' untranslated sequence, and nucleotides 1246 to 2243 are the 3' untranslated nucleotides. SEQ ID NO:2 represents the partial sequence of a soybean seed stearoyl-ACP desaturase cDNA. The first and last nucleotides (1 and 216 on clone 5) are read 5' to 3' and represent the

3' non-coding sequence. SEQ ID NO:3 represents the N-terminal sequence of the purified soybean seed stearoyl-ACP desaturase. SEQ ID NO:4 represents the degenerate coding sequence for amino acids 5 through 16 of SEQ ID NO:3. SEQ ID NO:5 represents a complementary mixture of degenerate oligonucleotides to SEQ ID NO:4.

The present invention is further defined in the following EXAMPLES, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these EXAMPLES, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these EXAMPLES, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

## EXAMPLE 1

## ISOLATION OF cDNA FOR SOYBEAN SEED STEAROYL-ACP DESATURASE

### PREPARATION OF [9,10-$^3$H]-STEAROYL-ACP

#### Purification of Acyl Carrier Protein (ACP) from E. coli

To frozen E. coli cell paste, (0.5 kg of 1/2 log phase growth of E. coli B grown on minimal media and obtained from Grain Processing Corp, Muscatine, IA) was added 50 mL of a solution 1 M in Tris, 1 M in glycine, and 0.25 M in EDTA. Ten mL of 1 M $MgCl_2$ was added and the suspension was thawed in a water bath at 50 °C. As the suspension approached 37 °C it was transferred to a 37 °C bath, made to 10 mM in 2-mercaptoethanol and 20 mg of DNAse and 50 mg of lysozyme were added. The suspension was stirred for 2 h, then sheared by three 20 second bursts in a Waring Blendor. The volume was adjusted to 1 L and the mixture was centrifuged at 24,000xg for 30 min. The resultant supernatant was centrifuged at 90,000xg for 2 h. The resultant high-speed pellet was saved for extraction of acyl-ACP synthase (see below) and the supernatant was adjusted to pH 6.1 by the addition of acetic acid. The extract was then made to 50% in 2-propanol by the slow addition of cold 2-propanol to the stirred solution at 0 °C. The resulting precipitate was allowed to settle for 2 h and then removed by centrifugation at 16,000xg. The resultant supernatant was adjusted to pH 6.8 with KOH and applied at 2 mL/min to a 4.4 x 12 cm column of DEAE-Sephacel® which had been equilibrated in 10 mM MES, pH 6.8. The column was washed with 10 mM MES, pH 6.8 and eluted with 1 L of a gradient of LiCl from 0 to 1.7 M in the same buffer. Twenty mL fractions were collected and the location of eluted ACP was determined by applying 10 µL of every second fraction to a lane of a native polyacrylamide (20% acrylamide) gel electrophoresis (PAGE). Fractions eluting at about 0.7 M LiCl contained nearly pure ACP and were combined, dialyzed overnight against water and then lyophilized.

#### Purification of Acyl-ACP Synthase

Membrane pellets resulting from the high-speed centrifugation described above were homogenized in 380 mL of 50 mM Tris-Cl, pH 8.0, and 0.5 M in NaCl and then centrifuged at 80,000xg for 90 min. The resultant supernatant was discarded and the pellets resuspended in 50 mM Tris-Cl, pH 8.0, to a protein concentration of 12 mg/mL. The membrane suspension was made to 2% in Triton X-100® and 10 mM in $MgCl_2$, and stirred at 0 °C for 20 min before centrifugation at 80,000xg for 90 min. The protein in the resultant supernatant was diluted to 5 mg/mL with 2% Triton X-100® in 50 mM Tris-Cl, pH 8.0 and, then, made to 5 mM ATP by the addition of solid ATP (disodium salt) along with an equimolar amount of $NaHCO_3$. The solution was warmed in a 55 °C bath until the internal temperature reached 53 °C and was then maintained at between 53 °C and 55 °C for 5 min. After 5 min the solution was rapidly cooled on ice and centrifuged at 15,000xg for 15 min. The supernatant from the heat treatment step was loaded directly onto a column of 7 mL Blue Sepharose® 4B which had been equilibrated in 50 mM Tris-Cl, pH 8.0, and 2% Triton X-100. The column was washed with 5 volumes of the loading buffer, then 5 volumes of 0.6 M NaCl in the same buffer and the activity was eluted with 0.5 M KSCN in the same buffer. Active fractions were assayed for the synthesis of acyl-ACP, as described below, combined, and bound to 3 mL settled-volume of hydroxylapatite equilibrated in 50 mM Tris-Cl, pH 8.0, 2% Triton X-100®. The hydroxylapatite was collected by centrifugation, washed twice with 20 mL of 50 mM Tris-Cl, pH 8.0, 2% Triton X-100®. The activity was eluted with two 5 mL washes of 0.5 M potassium phosphate, pH 7.5, 2% Triton X-100®. The first wash contained 66% of the activity and it was concentrated with a 30 kD membrane filtration concentrator (Amicon) to 1.5 mL.

Synthesis of [9,10-$^3$H]-Stearoyl-ACP

A solution of stearic acid in methanol (1 mM, 34.8 µL) was mixed with a solution of [9,10-$^3$H]stearate (Amersham) containing 31.6 µCi of $^3$H and dried in a glass vial. The ACP preparation described above (1.15 mL, 32 nmoles) was added along with 0.1 mL of 0.1 M ATP, 0.05 mL of 80 mM DTT, 0.1 mL of 8 M LiCl, and 0.2 mL of 13% Triton X-100® in 0.5 M Tris-Cl, pH 8.0, with 0.1 M $MgCl_2$. The reaction was mixed thoroughly and 0.3 mL of the acyl-ACP synthase preparation was added. After 1 h at 37 • C, a 10 µL aliquot was taken and dried on a small filter paper disc. The disc was washed extensively with chloroform:methanol:acetic acid (8:2:1, v:v:v) and radioactivity retained on the disc was taken as a measure of stearoyl-ACP. At 1 h about 67% of the ACP had been consumed and the reaction did not proceed further in the next 2 h. The reaction mix was diluted 1 to 4 with 20 mM Tris-Cl, pH 8.0, and applied to a 1 mL DEAE-Sephacel® column equilibrated in the same buffer. The column was washed in sequence with 5 mL of 20 mM Tris-Cl, pH 8.0, 5 mL of 80% 2-propanol in 20 mM Tris-Cl, pH 8.0, and eluted with 0.5 M LiCl in 20 mM Tris-Cl, pH 8.0. The column eluate was passed directly onto a 3 mL column of octyl-sepharose® CL-4B which was washed with 10 mL of 20 mM potassium phosphate, pH 6.8, and then eluted with 35% 2-propanol in 2 mM potassium phosphate, pH 6.8. The eluted volume (5.8 mL) contained 14.27 µCi of $^3$H (49% yield based on ACP). The eluted product was lyophilized and redissolved at a concentration of 24 µM [$^3$H]stearoyl-ACP at 0.9 mCi/µmole.

PREPARATION OF ALKYL-ACP AFFINITY COLUMN

Synthesis of N-hexadecyliodoacetamide

1-Hexadecylamine (3.67 mmole) was dissolved in 14.8 mL of $CH_2Cl_2$, cooled to 4 • C, and 2.83 mmoles of iodoacetic anhydride in 11.3 mL of $CH_2Cl_2$ was added dropwise to the stirred solution. The solution was warmed to room temperature and held for 2 h. The reaction mixture was diluted to about 50 mL with $CH_2Cl_2$ and washed 3 times (25 mL) with saturated sodium bicarbonate solution and then 2 times with water. The volume of the solution was reduced to about 5 mL under vacuum and passed through 25 mL of silica in diethyl ether. The eluate was reduced to an off-white powder under vacuum. This yielded 820 mg (2.03 mmoles) of the N-hexadecyliodoacetamide (71.8% yield). The 300 MHz $^1$H NMR spectra of the product was consistent with the expected structure.

Synthesis of N-Hexadecylacetamido-S-ACP

E. coli ACP prepared as above (10 mg in 2 mL of 50 mM Tris-Cl, pH 7.6) was treated at 37 • C with 50 mM DTT for 2 h. The solution was made to 10% TCA, held at 0 • C for 20 min and centrifuged to pellet. The resultant pellet was washed (2 x 2 mL) with 0.1 M citrate, pH 4.2 and redissolved in 3 mL of 50 mM potassium phosphate buffer. The pH of the ACP solution was adjusted to 7.5 with 1 M KOH and 3 mL of *N*-hexadecyliodoacetamide (3 mM in 2-propanol) was added. A slight precipitate of the *N*-hexadecyliodoacetamide was redissolved by warming the reaction mix to 45 • C. The mixture was held at 45 • C for 6 h. SDS-PAGE on 20% acrylamide PAGE gel showed approximately 80% conversion to an ACP species of intermediate mobility between the starting, reduced ACP and authentic palmitoyl-ACP. Excess *N*-hexadecyliodoacetamide was removed from the reaction mix by 4 extractions (3 mL) with $CH_2Cl_2$ with gentle mixing to avoid precipitation of the protein at the interface.

Coupling of N-Hexadecylacetamido-S-ACP to CNBr-activated Sepharose® 4B

Cyanogen bromide-activated Sepharose® 4B (Pharmacia, 2 g) was suspended in 1 mM HCl and extensively washed by filtration and resuspension in 1 mM HCl and finally one wash in 0.1 M $NaHCO_3$, pH 8.3. The N-hexadecylacetamido-S-ACP prepared above was diluted with an equal volume of 0.2 M $NaHCO_3$, pH 8.3. The filtered cyanogen bromide-activated Sepharose® 4B (about 5 mL) was added to the N-hexadecylacetamido-S-ACP solution, the mixture was made to a volume of 10 mL with the 0.1 M $NaHCO_3$, pH 8.3, and mixed by tumbling at room temperature for 6 h. Protein remaining in solution (Bradford assay) indicated approximately 85% binding. The gel suspension was collected by centrifugation, washed once with the 0.1 M $NaHCO_3$, pH 8.3, and resuspended in 0.1 M ethanolamine adjusted to pH 8.5 with HCl. The suspension was allowed to stand at 4 • C overnight and then washed by centrifugation and resuspension in 12 mL of 0.1 M acetate, pH 4.0, 0.5 M in NaCl and then 0.1 M $NaHCO_3$, pH 8.3, 0.5 M in NaCl. The alkyl-ACP Sepharose® 4B was packed into a 1 x 5.5 cm column and washed extensively with 20

mM bis-tris propane-Cl (BTP-Cl), pH 7.2, before use.

STEAROYL-ACP DESATURASE ASSAY

Stearoyl-ACP desaturase was assayed as described by McKeon et al. [(1982) J. Biol. Chem. 257:12141-12147] except for using [9,10-$^3$H]-stearoyl-ACP. Use of the tritiated substrate allowed assaying the enzyme activity by release of tritium as water, although the assay based on the tritium release underestimates desaturation by a factor of approximately 4 relative to that observed using $^{14}$C-stearoyl-ACP by the method of McKeon et al. [(1982) J. Biol. Chem 257:12141-12147], apparently because not all tritium is at carbons 9 and 10. Nevertheless, this modification makes the enzyme assay more sensitive, faster and more reliable. The reaction mix consisted of enzyme in 25 µL of 230 µg/mL bovine serum albumin (Sigma), 49 µg/mL catalase (Sigma), 0.75 mM NADPH, 7.25 µM spinach ferredoxin, and 0.35 µM spinach ferredoxin:NADPH$^+$ oxidoreductase, 50 mM Pipes, pH 6.0, and 1 µM [9,10-$^3$H]-stearoyl-ACP (0.9 mCi/µmole). All reagents, except for the Pipes buffer, labeled substrate and enzyme extract, were preincubated in a volume of 7.25 µL at pH 8.0 at room temperature for 10 min before adding 12.75 µL the Pipes buffer and labeled substrate stocks. The desaturase reaction was usually terminated after 5 min by the addition of 400 µL 10% trichloroacetic acid and 50 µL of 10 mg/mL bovine serum albumin. After 5 min on ice, the protein precipitate was removed by centrifugation at 13,000xg for 5 min. An aliquot of 425 µL was removed from the resultant supernatant and extracted twice with 2 mL of hexane. An aliquot of 375 µL of the aqueous phase following the second hexane extraction was added to 5 mL of ScintiVerse® Bio HP (Fisher) scintillation fluid and used to determine radioactivity released as tritium.

PURIFICATION OF SOYBEAN SEED STEAROYL-ACP DESATURASE

Developing soybean seeds, ca. 20-25 days after flowering, were harvested and stored at -80 • C until use. 300 g of the seeds were resuspended in 600 mL of 50 mM BTP-Cl, pH 7.2, and 5 mM dithiothreitol (DTT) in a Waring Blendor. The seeds were allowed to thaw for a few minutes at room temperature to 4 • C and all of the purification steps were carried out at 4 • C unless otherwise noted. The seeds were homogenized in the blendor three times for 30 s each and the homogenate was centrifuged at 14,000xg for 20 min. The resultant supernatant was centrifuged at 100,000xg for 1 h. The resultant high-speed supernatant was applied, at a flow-rate of 5 mL/min to a 2.5 x 20 cm Blue Sepharose® column equilibrated in 10 mM BTP-Cl, pH 7.2, 0.5 mM DTT. Following a wash with 2 column volumes of 10 mM BTP-Cl, pH 7.2, 0.5 mM DTT, the bound proteins were eluted in the same buffer containing 1 M NaCl. The eluting protein peak, which was detected by absorbance at 280 nm, was collected and precipitated with 80% ammonium sulfate. Following collection of the precipitate by centrifugation at 10,000xg for 20 min, its resuspension in 10 mM potassium phosphate, pH 7.2, 0.5 mM DTT, overnight dialysis in the same buffer precipitate, and clarification through a 0.45 micron filter, it was applied to a 10 mm x 25 cm Wide-pore™ PEI ($NH_2$) anion-exchange column (Baker) at 3 mL/min thoroughly equilibrated in buffer A (10 mM potassium phosphate, pH 7.2). After washing the column in buffer A until no protein was eluted, the column was subjected to elution by a gradient from buffer A at 0 min to 0.25 M potassium phosphate (pH 7.2) at 66 min at a flow rate of 3 mL/min. Three mL fractions were collected. The desaturase activity eluted in fractions 17-25 (the activity peak eluted at ca. 50 mM potassium phosphate). The pooled fractions were made to 60 mL with buffer A and applied at 1 mL/min to a 1 x 5.5 cm alkyl-ACP column equilibrated in buffer A containing 0.5 mM DTT. After washing the bound protein with the start buffer until no protein was eluted, the bound protein was eluted by a gradient from buffer A containing 0.5 mM DTT at 0 min to 0.5 M potassium phosphate, pH 7.2, 0.5 mM DTT at 60 min and 1 M potassium phosphate, pH 7.2, 0.5 mM DTT. Four mL fractions were collected. Fractions 15-23, which contained the enzyme with the highest specific activity, were pooled and concentrated to 3 mL by a 30 kD Centricon® concentrator (Millipore) and desalted in a small column of G-25 Sephadex® equilibrated with 25 mM bis-Tris-Cl, pH 6.7. The desalted sample was applied at 1 mL/min to a chromatofocussing Mono P HR 5/20 (Pharmacia) column equilibrated with 25 mM bis-Tris-Cl, pH 6.7, washed with a column volume of the same buffer, and eluted with 1:10 dilution of Polybuffer 74 (Pharmacia) made to pH 5.0 with HCl. Desaturase activity eluted in two peaks: one in fraction 30 corresponding to a pI of ca. 6.0 and the other in fraction 35, corresponding to a pI of ca. 5.7. The protein in the two peaks were essentially composed of ca. 38 kD polypeptide. The first peak had a higher enzyme specific activity and was used for further characterization as well as for further purification on reverse-phase chromatography.

Mono P fractions containing the first peak of enzyme activity were pooled and applied to a $C_4$ reverse-phase HPLC column (Vydac) equilibrated with buffer A (5% acetonitrile, 0.1% trifluoroacetic acid) and

eluted at 0.1 mL/min with a gradient of 25% buffer B (100% acetonitrile, 0.1% trifluoroacetic acid) and 75% buffer A at 10 min to 50% buffer B and 50% buffer A at 72.5 min. A single major peak eluted at 41.5% buffer B that also ran as a ca. 38 kD protein based on SDS-PAGE. The protein in the peak fraction was used to determine the N-terminal amino acid sequence on a Applied Biosystems 470A Gas Phase Sequencer. The PTH amino acids were analysed on Applied Biosystems 120 PTH Amino Acid Analyzer.

The N-terminal sequence of the ca. 38 kD polypeptide was determined through 16 residues and is shown in SEQ ID NO:3.

CLONING OF SOYBEAN SEED STEAROYL-ACP DESATURASE cDNA

Based on the N-terminal amino acid sequence of the purified soybean seed stearoyl-ACP desaturase (SEQ ID NO:3), amino acids 5 through 16, which are represented by the degenerate coding sequence, SEQ ID NO:4, was chosen to design the complementary mixture of degenerate oligonucleotides (SEQ ID NO:5).

The design took into account the codon bias in representative soybean seed genes encoding Bowman-Birk protease inhibitor [Hammond et al. (1984) J. Biol. Chem. 259:9883-9890], glycinin subunit A-2B-1a (Utsumi et al. (1987) Agric. Biol. Chem. 51:3267-3273], lectin (le-1)[Vodkin et al. (1983) Cell 34:1023-1031], and lipoxygenase-1 [Shibata et al. (1987) J. Biol. Chem. 262:10080-10085]. Five deoxyinosines were used at selected positions of ambiguity.

A cDNA library was made as follows: Soybean embryos (ca. 50 mg fresh weight each) were removed from the pods and frozen in liquid nitrogen. The frozen embryos were ground to a fine powder in the presence of liquid nitrogen and then extracted by Polytron homogenization and fractionated to enrich for total RNA by the method of Chirgwin et al. [Biochemistry (1979) 18:5294-5299]. The nucleic acid fraction was enriched for poly $A^+$ RNA by passing total RNA through an oligo-dT cellulose column and eluting the poly $A^+$ RNA by salt as described by Goodman et al. [(1979) Meth. Enzymol. 68:75-90]. cDNA was synthesized from the purified poly $A^+$ RNA using cDNA Synthesis System (Bethesda Research Laboratory) and the manufacturer's instructions. The resultant double-stranded DNA was methylated by DNA methylase (Promega) prior to filling-in its ends with T4 DNA polymerase (Bethesda Research Laboratory) and blunt-end ligating to phosphorylated Eco RI linkers using T4 DNA ligase (Pharmacia). The double-stranded DNA was digested with Eco RI enzyme, separated from excess linkers by passing through a gel filtration column (Sepharose CL-4B), and ligated to Lambda ZAP vector (Stratagene) as per manufacturer's instructions. Ligated DNA was packaged into phage using Gigapack packaging extract (Stratagene) according to manufacturer's instructions. The resultant cDNA library was amplified as per Stratagene's instructions and stored at -80 °C.

Following the instructions in Lambda ZAP Cloning Kit Manual (Stratagene), the cDNA phage library was used to infect E. coli BB4 cells and plated to yield ca. 80,000 plaques per petri plate (150 mm diameter). Duplicate lifts of the plates were made onto nitrocellulose filters (Schleicher & Schuell). Duplicate lifts from five plates were prehybridized in 25 mL of Hybridization buffer consisting of 6X SSC (0.9 M NaCl, 0.09 M sodium citrate, pH 7.0), 5X Denhardt's [0.5 g Ficoll (Type 400, Pharmacia), 0.5 g polyvinylpyrrolidone, 0.5 g bovine serum albumin (Fraction V; Sigma)], 1 mM EDTA, 1% SDS, and 100 ug/mL denatured salmon sperm DNA (Sigma Chemical Co.) at 45 °C for 10 h. Ten pmol of the hybridization probe (see above) were end-labeled in a 52.5 uL reaction mixture containing 50 mM Tris-Cl, pH 7.5, 10 mM $MgCl_2$, 0.1 mM spermidine-HCl (pH 7.0), 1 mM EDTA (pH 7.0), 5 mM DDT, 200 uCi (66.7 pmoles) of gamma-labeled $AT^{32}P$ (New England Nuclear) and 25 units of T4 polynucleotide kinase (New England Biolabs). After incubation at 37 °C for 45 min, the reaction was terminated by heating at 68 °C for 10 min. Labeled probe was separated from unincorporated $AT^{32}P$ by passing the reaction through a Quick-Spin™ (G-25 Sephadex®) column (Boehringer Mannheim Biochemicals). The purified labeled probe ($1.2 \times 10^7$ dpm/pmole) was added to the prehybridized filters, following their transfer to 10 mL of fresh Hybridization buffer. Following incubation of the filters in the presence of the probe for 16 h in a shaker at 48 °C, the filters were washed in 200 mL of Wash buffer (6X SSC, 0.1% SDS) five times for 5 min each at room temperature, and then once at 48 °C for 5 min. The washed filters were air dried and subjected to autoradiography on Kodak XAR-2 film in the presence of intensifying screens (Lightening Plus, DuPont Cronex®) at -80 °C overnight. Six positively-hybridizing plaques were subjected to plaque purification as described in Sambrook et al. [(1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press]. Following the Lambda ZAP Cloning Kit Instruction Manual (Stratagene), sequences of the pBluescript vector, including the cDNA inserts, from each of six purified phages were excised in the presence of a helper phage and the resultant phagemids were used to infect E. coli XL-1 Blue cells resulting in double-stranded plasmids, pDS1 to pDS6. The restriction maps of all six plasmids, though not identical, showed a common 0.7 kb Bgl II fragment found in the desaturase gene (see below).

DNA from plasmids pDS1-pDS6 were made by the alkaline lysis miniprep procedure described in Sambrook et al. [(1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press]. The alkali-denatured double-stranded DNAs were sequenced using Sequenase® T7 DNA polymerase (US Biochemical Corp.) and the manufacturer's instructions. The sequence of the cDNA insert in plasmid pDS1 is shown in SEQ ID NO:1.

## EXAMPLE 2

## EXPRESSION OF SOYBEAN SEED STEAROYL-ACP DESATURASE IN E. COLI

### Construction of Glutathione-S-Transferase: Stearoyl-ACP Desaturase Fusion Protein

Plasmid pDS1 was linearized with Hind III enzyme, its ends filled-in with Klenow fragment (Bethesda Research Laboratory) in the presence of 50 µM each of all four deoxynucleotide triphosphates as per manufacturer's instructions, and extracted with phenol:chloroform (1:1). Phosphorylated Eco RI linkers (New England Biolabs) were ligated to the DNA using T4 DNA ligase (New England Biolabs). Following partial digestion with Bgl II enzyme and complete digestion with excess Eco RI enzyme, the DNA was run on an agarose gel and stained with ethidium bromide. The 2.1 kb DNA fragment resulting from a partial Bgl II and Eco RI digestion was cut out of the gel, purified using USBioclean™ (US Biochemicals), and ligated to Bam HI and Eco RI cleaved vector pGEX2T [Pharmacia; see Smith et al. (1988) Gene 67:31] using T4 DNA ligase (New England Biolabs). The ligated mixture of DNAs were used to transform E. coli XL-1 blue cells (Stratagene). Transformants were picked as ampicillin-resistant cells and the plasmid DNA from several transformants analyzed by digestion with Bam HI and Eco RI double restriction digest, as described by Sambrook et al. [(1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press]. Plasmid DNA from one transformant, called pGEXB, showed the restriction pattern expected from the correct fusion. The double-stranded plasmid pGEXB was purified and sequenced to confirm the correct fusion by the Sequenase kit (US Biochemical Corp.). The fusion protein consists of a 28 kD glutathione-S-transferase protein fused at its C-terminus to the desaturase precursor protein at Ser at residue -10 from the N-terminus of the mature enzyme (Arg, +1) (SEQ ID NO:1). Thus, it includes ten amino acids from the transit peptide sequence in addition to the mature protein.

### Inducible Expression of the Glutathione-S-Transferase-Stearoyl-ACP Desaturase Fusion Protein

Five mL precultures of plasmids pGEXB and pGEX2T, which were grown overnight at 37°C in LB medium [Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press] containing 100 ug/mL ampicillin, were diluted 1:10 in fresh LB medium containing 100 µg/mL ampicillin and continued to grow on a shaker at 37°C for another 90 min before adding isopropylthio-β-D-galactoside and ferric chloride to final concentrations of 0.3 mM and 50 µM, respectively. After an additional 3 h on a shaker at 37°C, the cultures were harvested by centrifugation at 4,000xg for 10 min at 4°C. The cells were resuspended in one-tenth of the culture volume of freshly-made and ice-cold Extraction buffer (20 mM sodium phosphate, pH 8.0, 150 mM NaCl, 5 mM EDTA and 0.2 mM phenylmethylsulfonyl fluoride) and re-centrifuged as above. The resultant cells were resuspended in 1/50 vol of the culture in Extraction buffer and sonicated for three ten-second bursts. The sonicated extracts were made to 1% in Triton X-100 and centrifuged at 8,000xg for 1 min in Eppendorf Micro Centrifuge (Brinkmann Instruments) to remove the cellular debris. The supernatant was poured into a fresh tube and used for enzyme assays, SDS-PAGE analysis and purification of the fusion protein.

Five µL aliquots of the extracts were assayed for stearoyl-ACP desaturase activity in a 1 min reaction, as described in Example I. The activities [net pmol of stearoyl-ACP desaturated per min per mL of extract; the blank (no desaturase enzyme) activity was 15 pmol/min/ml] are shown below:

| Reaction mixture | Net pmol/min/mL |
|---|---|
| E. coli (pGEX2T) | 0 |
| E. coli (pGEXB) | 399 |
| E. coli (pGEXB) - NADPH | 0 |
| E. coli (pGEXB) - ferredoxin | 0 |
| E. coli (pGEXB) - ferredoxin-NADPH reductase | 3 |

These results show that the desaturase enzyme activity is present in the extract of E. coli cells containing pGEXB but not in that of cells containing the control plasmid pGEX2T. Furthermore, this activity was dependent on an exogenous electron donor.

Proteins in extracts of E. coli cells harboring plasmids pGEX2T or pGEXB were resolved by SDS-PAGE, transferred onto Immobilon®-P (Millipore) and cross-reacted with mouse antibody made against purified soybean stearoyl-ACP desaturase, as described by Sambrook et al. [(1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press]. The resultant Western blot showed that pGEXB encodes for ca. 64 kD GST-stearoyl-ACP desaturase fusion polypeptide, although some lower molecular-weight cross-reacting polypeptides can also be observed, which may represent either a degradation or incomplete synthesis of the fusion protein. It is not known whether the GST-desaturase fusion protein is enzymatically active, since the activity observed may be due to the incomplete fusion by the peptides seen here. The fusion polypeptide was not present in extracts of cells harboring the control plasmid (pGEX2T) nor in extracts of cells harboring pGEXB that were not induced by isopropylthio-β-D-galactoside.

Purification of the Glutathione-S-Transferase-Stearoyl-ACP Desaturase Fusion Protein

The GST-desaturase fusion protein was purified in a one step glutathione-agarose affinity chromatography under non-denaturing conditions, following the procedure of Smith et al. [Gene (1988) 67:31]. For this, the bacterial cell extract was mixed with 1 mL glutathione-agarose (sulfur-linkage, Sigma), equilibrated with 20 mM sodium phosphate, pH 8.0, 150 mM NaCl, for 10 min at room temperature. The beads were collected by centrifugation at 1000xg for 1 min, and washed three times with 1 mL of 20 mM sodium phosphate, pH 8.0, 150 mM NaCl (each time the beads were collected by centrifugation as described above). The fusion protein was eluted with 5 mM reduced glutathione (Sigma) in 50 mM Tris-Cl, pH 8.0. The proteins in the eluted fraction were analyzed by SDS-PAGE and consisted of mostly pure ca. 64 kD GST-desaturase polypeptide, 28 kD GST and a trace of ca. 38 kD desaturase polypeptide. The fusion polypeptide was cleaved in the presence of thrombin, as described by Smith et al. [Gene (1988) 67:31].

Construction of β-Galactosidase-Stearoyl-ACP Desaturase Fusion Protein

Plasmid pDS1 DNA was digested with Ssp I and Pvu I enzymes and the digested DNA fragments were resolved by electrophoresis in agarose. The blunt-ended 2.3 kb Ssp I fragment was cut out of the agarose (Pvu I cleaves a contaminating 2.3 kb Ssp I fragment), purified by USBioclean™ (US Biochemical Corp.), and ligated to vector plasmid pBluescript SK (-) (Stratagene) that had previously been filled-in with Klenow fragment (Bethesda Research Laboratory) following linearization with Not I enzyme. The ligated DNAs were transformed into competent E. coli XL-1 blue cells. Plasmid DNA from several ampicillin-resistant transformants were analysed by restriction digestion. One plasmid, called pNS2, showed the expected physical map. This plasmid is expected to encode a ca. 42 kD fusion protein consisting of 4 kD N-terminal of β-galactosidase fused at its C-terminus to isoleucine at residue +10 from the N-terminus of the mature desaturase protein (Arg, +1) (SEQ ID NO:1). Thus, it includes all but the first 10 amino acids of the mature protein. Nucleotide sequencing has not been performed on pNS2 to confirm correct fusion.

Five mL of preculture of E. coli cells harboring plasmid pNS2 grown overnight in LB medium containing 100 µg/mL ampicillin was added to 50 mL of fresh LB medium with 100 µg/mL ampicillin. After additional 1 h of growth at 37°C in a shaker, isopropylthio-β-D-galactoside and ferric chloride were added to final concentrations of 0.3 mM and 50 µM, respectively. After another 2 h on a shaker at 37°C, the culture was harvested by centrifugation at 4,000xg for 10 min at 4°C. The cells were resuspended in 1 mL of freshly-made and ice-cold TEP buffer (100 mM Tris-Cl, pH 7.5, 10 mM EDTA and 0.1 mM phenylmethylsulfonyl fluoride) and recentrifuged as above. The cells were resuspended in 1 mL of TEP buffer and sonicated for three ten-second bursts. The sonicates were made to 1% in Triton X-100, allowed to stand in ice for 5 min, and centrifuged at 8,000xg for 1 min in an Eppendorf Micro Centrifuge (Brinkmann Instruments) to remove the cellular debris. The supernatant was poured into a fresh tube and used for enzyme assays and SDS-PAGE analysis.

A 1 µL aliquot of the extract of E. coli cells containing plasmid pNS2 was assayed for stearoyl-ACP desaturase activity in a 5 min reaction, as described above. The extract showed activity of 288 pmol of stearoyl-ACP desaturated per min per ml of the extract [The blank (no desaturase enzyme) activity was 15 pmol/min/mL].

Proteins in the extract of E. coli cells harboring plasmids pNS2 were resolved by SDS-PAGE, transferred onto Immobilon®-P (Millipore) and cross-reacted with mouse antibody made against purified soybean stearoyl-ACP desaturase, as described in Sambrook et al. [(1989) Molecular Cloning: A Laboratory

Manual, 2nd Ed. Cold Spring Harbor Laboratory Press]. The resultant Western blot showed that pNS2 encodes for ca. 42 kD β-galactosidase-stearoyl-ACP desaturase fusion polypeptide.

EXAMPLE 3

USE OF SOYBEAN SEED STEAROYL-ACP DESATURASE SEQUENCE IN PLASMID pDS1 AS A RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP) MARKER

Plasmid pDS1 was linearized by digestion with restriction enzyme Eco RI in standard conditions as described in Sambrook et al. [(1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press] and labeled with $^{32}P$ using a Random Priming Kit from Bethesda Research Laboratories under conditions recommended by the manufacturer. The resulting radioactive probe was used to probe a Southern blot [Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press] containing genomic DNA from soybean [Glycine max (cultivar Bonus) and Glycine soja - (PI81762)], digested with one of several restriction enzymes. After hybridization and washes under standard conditions [Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory Press] autoradiograms were obtained and different patterns of hybridization (polymorphisms) were identified in digests performed with restriction enzymes Pst 1 and Eco RI. The same probe was then used to map the polymorphic pDS1 loci on the soybean genome, essentially as described by Helentjaris et al. [(1986) Theor. Appl. Genet. 72:761-769]. Plasmid pDS1 probe was applied, as described above, to Southern blots of Eco RI or Pst I digested genomic DNAs isolated from 68 F2 progeny plants resulting from a G. max Bonus x G. soja PI81762 cross. The bands on the autoradiograms were interpreted as resulting from the inheritance of either paternal (Bonus) or maternal (PI81762) pattern, or both (a heterozygote). The resulting data were subjected to genetic analysis using the computer program Mapmaker [Lander et al., (1987) Genomics 1: 174-181]. In conjunction with previously obtained data for 436 anonymous RFLP markers in soybean [Tingey et al. (1990) J. Cell. Biochem., Supplement 14E p. 291, abstract R153], we were able to position four genetic loci corresponding to the pDS1 probe on the soybean genetic map. This information will be useful in soybean breeding targeted towards developing lines with altered saturate levels, especially for the high stearic acid mutant phenotype, since these recessive traits are most likely be due to loss of seed stearoyl-ACP desaturase enzyme.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Hitz, William D.
Yadav, Narendra S

(ii) TITLE OF THE INVENTION: Nucleotide Sequence of SoybeanStearoyl-ACP Desaturase cDNA

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: E. I. du Pont de Nemours and Company
(B) STREET: 1007 Market Street
(C) CITY: Wilmington
(D) STATE: Delaware
(E) COUNTRY: USA
(F) ZIP: 19898

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: DISKETTE, 3.50 inch, 1.0 MB
(B) COMPUTER: Apple Macintosh
(C) OPERATING SYSTEM:
(D) SOFTWARE:

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 07/529,049
(B) FILING DATE: 25-MAY-1990
(C) CLASSIFICATION:

(vii) ATTORNEY/AGENT INFORMATION;

(A) NAME: Bruce W. Morrissey

(B) REGISTRATION NUMBER: 30,663

(C) REFERENCE/DOCKET NUMBER: BB-1022

(viii) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (302) 892-4927

(B) TELEFAX: (302) 892-7949

(C) TELEX: 835420

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:2243 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: No

(iv) ANTISENSE: No

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Glycine max

(B) STRAIN: Cultivar Wye

(D) DEVELOPMENTAL STAGE: Developing seeds

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA to mRNA

(B) CLONE: pDS1

(ix) FEATURE:

(A) NAME/KEY:

(i) 5' non-coding sequence

(ii) Putative translation initiation codon

(iii) Putative transit peptide coding sequence

(iv) Mature protein coding sequence

(v) Translation termination codon

(vi) 3' non-coding sequence

(B) LOCATION:

(i) nucleotides 1 through 69

(ii) nucleotides 70 through 72

(iii) nucleotides 70 through 165

(iv) nucleotides 166 through 1242

(v) nucleotides 1243 through 1245

(vi) nucleotides 1246 through 2243

(C) IDENTIFICATION METHOD:

(i) deduced by proximity to ii) below

(ii) similarity of the context of the methionine codon in the open reading frame to translation initiation codons of other plastid transit peptides

(iii) deduced by proximity to ii) above and iv) below

(iv) experimental determination of N-terminal amino acid sequence and subunit size of purified soybean seed stearoyl-ACP desaturase

(v) The translation termination codon ends the open reading frame for a protein of the expected size

(vi) established by proximity to v) above

(D) OTHER INFORMATION: Extracts of E. coli expressing the mature protein as a fusion protein show stearoyl-ACP desaturase activity and produce a protein that cross-reacts to stearoyl-ACP desaturase antibody

(x) PUBLICATION INFORMATION: Sequence not published.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
          CTTCTACATT ACTCTCTCTT CTCCTAAAAA TTTCTAATGC          40

TTCCATTGCT TCATCTGACT CACTCATCA ATG GCT CTG AGA CTG AAC CCT   90
                                Met Ala Leu Arg Leu Asn Pro
                                -32     -30

ATC CCC ACC CAA ACC TTC TCC CTC CCC CAA ATG CCC AGC CTC AGA  135
Ile Pro Thr Gln Thr Phe Ser Leu Pro Gln Met Pro Ser Leu Arg
-25                 -20                 -15

TCT CCC CGC TTC CGC ATG GCT TCC ACC CTC CGC TCC GGT TCC AAA  180
Ser Pro Arg Phe Arg Met Ala Ser Thr Leu Arg Ser Gly Ser Lys
-10                 -5                   1                   5
```

```
GAG GTT GAA AAT ATT AAG AAG CCA TTC ACT CCT CCC AGA GAA GTG  225
Glu Val Glu Asn Ile Lys Lys Pro Phe Thr Pro Pro Arg Glu Val
                10                  15                  20

CAT GTT CAA GTA ACC CAC TCT ATG CCT CCC CAG AAG ATT GAG ATT  270
His Val Gln Val Thr His Ser Met Pro Pro Gln Lys Ile Glu Ile
                25                  30                  35

TTC AAA TCT TTG GAG GAT TGG GCT GAC CAG AAC ATC TTG ACT CAT  315
Phe Lys Ser Leu Glu Asp Trp Ala Asp Gln Asn Ile Leu Thr His
                40                  45                  50

CTT AAA CCT GTA GAA AAA TGT TGG CAA CCA CAG GAT TTT TTA CCC  360
Leu Lys Pro Val Glu Lys Cys Trp Gln Pro Gln Asp Phe Leu Pro
                55                  60                  65

GAC CCC TCC TCA GAT GGA TTT GAA GAG CAA GTG AAG GAA CTG AGA  405
Asp Pro Ser Ser Asp Gly Phe Glu Glu Gln Val Lys Glu Leu Arg
                70                  75                  80

GAG AGA GCA AAG GAG ATT CCA GAT GAT TAC TTT GTT GTT CTT GTC  450
Glu Arg Ala Lys Glu Ile Pro Asp Asp Tyr Phe Val Val Leu Val
                85                  90                  95

GGA GAC ATG ATC ACA GAG GAA GCT CTG CCT ACT TAC CAA ACT ATG  495
Gly Asp Met Ile Thr Glu Glu Ala Leu Pro Thr Tyr Gln Thr Met
                95                  100                 110

TTA AAT ACT TTG GAT GGA GTT CGT GAT GAA ACA GGT GCC AGC CTT  540
Leu Asn Thr Leu Asp Gly Val Arg Asp Glu Thr Gly Ala Ser Leu
                115                 120                 125

ACT TCC TGG GCA ATT TGG ACA AGG GCA TGG ACT GCT GAA GAA AAC  585
Thr Ser Trp Ala Ile Trp Thr Arg Ala Trp Thr Ala Glu Glu Asn
                130                 135                 140

AGA CAC GGT GAT CTT CTT AAC AAA TAT CTG TAC TTG AGT GGA CGA  630
Arg His Gly Asp Leu Leu Asn Lys Tyr Leu Tyr Leu Ser Gly Arg
                145                 150                 155

GTT GAC ATG AAA CAA ATT GAG AAG ACA ATT CAG TAC CTT ATT GGG  675
Val Asp Met Lys Gln Ile Glu Lys Thr Ile Gln Tyr Leu Ile Gly
                160                 165                 170

TCT GGG ATG GAT CCT CGG ACC GAG AAC AGC CCC TAC CTT GGT TTC  720
Ser Gly Met Asp Pro Arg Thr Glu Asn Ser Pro Tyr Leu Gly Phe
                175                 180                 185

ATT TAC ACT TCA TTT CAA GAG AGG GCA ACC TTC ATA TCC CAC GGA  765
Ile Tyr Thr Ser Phe Gln Glu Arg Ala Thr Phe Ile Ser His Gly
                190                 195                 200

AAC ACG GCC AGG CTT GCG AAG GAG CAT GGT GAC ATA AAA TTG GCA  810
Asn Thr Ala Arg Leu Ala Lys Glu His Gly Asp Ile Lys Leu Ala
                205                 210                 215
```

```
CAG ATC TGC GGC ATG ATT GCC TCA GAT GAG AAG CGC CAC GAG ACT  855
Gln Ile Cys Gly Met Ile Ala Ser Asp Glu Lys Arg His Glu Thr
                220                 225                 230

GCA TAC ACA AAG ATA GTG GAA AAG CTG TTT GAG GTT GAT CCT GAT  900
Ala Tyr Thr Lys Ile Val Glu Lys Leu Phe Glu Val Asp Pro Asp
                235                 240                 245

GGT ACA GTT ATG GCA TTT GCC GAC ATG ATG AGG AAG AAG ATT GCT  945
Gly Thr Val Met Ala Phe Ala Asp Met Met Arg Lys Lys Ile Ala
                250                 255                 260

ATG CCA GCA CAC CTT ATG TAT GAC GGC CGC GAC GAC AAC CTG TTT  990
Met Pro Ala His Leu Met Tyr Asp Gly Arg Asp Asp Asn Leu Phe
                265                 270                 275

GAT AAC TAC TCT GCC GTC GCG CAG CGC ATT GGG GTC TAC ACT GCA 1035
Asp Asn Tyr Ser Ala Val Ala Gln Arg Ile Gly Val Tyr Thr Ala
                280                 285                 290

AAG GAC TAT GCT GAC ATA CTC GAA TTT CTG GTG GGG AGG TGG AAG 1080
Lys Asp Tyr Ala Asp Ile Leu Glu Phe Leu Val Gly Arg Trp Lys
                295                 300                 305

GTG GAG CAG CTA ACC GGA CTT TCA GGT GAG GGA AGA AAG GCT CAG 1125
Val Glu Gln Leu Thr Gly Leu Ser Gly Glu Gly Arg Lys Ala Gln
                310                 315                 320

GAA TAC GTT TGT GGG CTG CCA CCA AGA ATC AGA AGG TTG GAG GAG 1170
Glu Tyr Val Cys Gly Leu Pro Pro Arg Ile Arg Arg Leu Glu Glu
                325                 330                 335

AGA GCT CAA GCA AGA GGC AAG GAG TCG TCA ACA CTT AAA TTC AGT 1215
Arg Ala Gln Ala Arg Gly Lys Glu Ser Ser Thr Leu Lys Phe Ser
                340                 345                 350

TGG ATT CAT GAC AGG GAA GTA CTA CTC TAAATGCT TGCACCAAGG     1260
Trp Ile His Asp Arg Glu Val Leu Leu
                355             359

GAGGAGCATG GTGAATCTTC CAGCAATACC ATTCTGAGAA ATGTTGAATA      1310

GTTGAAAATT CAGTTTGTCA TTTTTATCTT TTTTTTCTCC TGTTTTTTGG      1360

TCTTATGTTA TATGCCACTG TAAGGTGAAA CAGTTGTTCT TGCATGGTTC      1410

GCAAGTTAAG CAGTTAGGGG CAGCTGTAGT ATTAGAAATG CTATTTTTTG      1460

TTTCCCTTTT CTGTGGTAGT GATGTCTGTG GAAGTATAAG TAAACGTTTT      1510

TTTTTTCTC TGGCAATTTTG ATGATAAAGA AAATTTAGTT CTAAAAACCG      1560

TCGCACCTTC CCTGAGGCTT CTCTTGTCTG TCGCGAGTGA CCATGGTGAG      1610

GGTTAGTGTG CTGAACGATG CTCTGAAGAG CATGTACAAT GCTGAGAAAA      1660

GGGGAAAGCG CCAAGTCATG ATTCGGCCAT CCTCCAAAGT CATTATCAAA      1710
```

```
TTCCTTTTGG TGATGCAGAA GCACGGATAC ATTGGAGAGT TTGAGTATGT      1760

TGATGACCAC AGGGCTGGTA AAATCGTGGT TGAATTGAAC GGTAGACTGA      1810

ACAAGTGTGG GGTTATTAGT CCCCGTTTTG ATGTCGGCGT CAAAGAGATT      1860

GAAGGTTGGA CTGCTAGGCT TCTCCCCTCA AGACAGTTTG GGTATATTGT      1910

ATTGACTACC TCTGCCGGCA TCATGGATCA CGAAGAAGCT AGGAGAAAAA      1960

ATGTTGGTGG TAAGGTACTG GGTTTCTTCT ACTAGAGTTT AATTTCGATT      2010

AAGAGGATGT CAGGAATTTC AATTGAGATT CATGGATTGT AATGGAGGAT      2060

ATGCTAGGCC CCTAGTAATA TCAAGCATAG CAGGAGCTGT TTTGTGATGT      2110

TCCTTATTTT GTTTGCAAAA CCAAGTTGGT AACTATAACT TTTATTTTCT      2160

TTTATCATTA TTTTTCTTTA TACCAAAATG TACTGGCCAA GTTGTTTTAA      2210

ACAGTGAGAA CTTTGATTAG AAAAAAAAAA AAA                        2243
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH:216 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: No

(iv) ANTISENSE: No

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Glycine max
- (B) STRAIN: Cultivar Wye
- (D) DEVELOPMENTAL STAGE: Developing seeds

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: cDNA to mRNA
(B) CLONE: pDS4a

(ix) FEATURE:
(A) NAME/KEY:3' non-coding sequence
(B) LOCATION: nucleotides 1 through 216
(C) IDENTIFICATION METHOD: Homology of clones pDS4a and pDS1 and similarity of sequence in SEQ ID NO:1 to 3' non-coding sequence in SEQ ID NO:1

(x) PUBLICATION INFORMATION: Sequence not published.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
GAAATGTTGA ATAGTTGAAA ATTCAGTTTG TCATTTTTAT CTTTTATTTT  50
TTCTCCTTTT TTGGTCTTTG TTATATGTCA CTGTAAGGTG AAGCAGTTGT 100
TCTTGCATGG TTCGCAAGTT AAGCAGTTAG GGGCAGCTGT AGTATTAGAA 150
ATGGTATTTT TTTTTTGTT TTCGCTTTTC TCTGTGGTAG TGATGTCTGT 200
CGAAGTATAA GTAAAC                                     216
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: No

(v) FRAGMENT TYPE: N-terminal fragment

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Glycine max
(B) STRAIN: Cultivar Wye
(C) DEVELOPMENTAL STAGE: Developing seeds

(ix) FEATURE:
(A) NAME/KEY: N-terminal sequence
(B) LOCATION: 1 through 16 amino acid residues
(C) IDENTIFICATION METHOD: N-terminal amino acid sequencing

(x) PUBLICATION INFORMATION: Sequence not published

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

| Arg | Ser | Gly | Ser | Lys | Glu | Val | Glu | Asn | Ile | Lys | Lys | Pro | Phe | Thr | Pro |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | 5 | | | | | 10 | | | | | | 15 |

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid: mixture of oligonucleotides

(iii) HYPOTHETICAL: Yes

(ix) FEATURE:

(A) NAME/KEY: Coding sequence

(B) LOCATION: 1 through 36 bases

(x) PUBLICATION INFORMATION : Sequence not published

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
AAR GAR GTN GAR AAY ATH AAR AAR CCN TTY ACN CCN  3
Lys Glu Val Glu Asn Ile Lys Lys Pro Phe Thr Pro
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:35 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid: mixture of synthetic oligonucleotides

(ix) FEATURE:

(C) OTHER INFORMATION: N at positions 3,6,9, and 27 is deoxyinosine.

(x) PUBLICATION INFORMATION: Sequence not published

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGNGTNAANG GCTTCTTRAT RTTYTCNACN TCCTT 35

## Claims

1. An isolated nucleic acid fragment comprising a nucleotide sequence encoding the soybean seed stearoyl-ACP desaturase corresponding to the nucleotides 1 to 2243 in SEQ ID NO:1, or any soybean nucleic acid fragment substantially homologous therewith encoding a functional stearoyl-ACP desaturase.

2. An isolated nucleic acid fragment of Claim 1 wherein said nucleotide sequence encodes the soybean seed stearoyl-ACP desaturase precursor corresponding to nucleotides 70-1245 in SEQ ID NO:1, or any soybean nucleic acid fragment substantially homologous therewith encoding a functional stearoyl-ACP desaturase precursor.

3. A nucleic acid fragment of Claim 2, wherein the said nucleotide sequence encodes the mature soybean seed stearoyl-ACP desaturase enzyme, corresponding to nucleotides 166 to 1245 in SEQ ID NO:1.

4. A chimeric gene capable of transforming a soybean plant cell comprising a nucleic acid fragment of Claim 1 operably linked to suitable regulatory sequences producing antisense inhibition of soybean seed stearoyl-ACP desaturase in the seed.

5. A chimeric gene capable of transforming a plant cell of an oil-producing species comprising a nucleic acid fragment of Claim 2 operably linked to suitable regulatory sequences resulting in overexpression of said soybean seed stearoyl-ACP desaturase in the plastid of said plant cell.

6. A chimeric gene capable of transforming a plant cell of an oil-producing species comprising a nucleic acid fragment of Claim 3 operably linked to suitable regulatory sequences resulting in the expression of said mature soybean seed stearoyl-ACP desaturase enzyme in the cytoplasm of said plant cell.

7. A method of producing soybean seed oil containing higher-than-normal levels of stearic acid comprising:
   (a) transforming a soybean plant cell with a chimeric gene of Claim 4,
   (b) growing fertile soybean plants from said transformed soybean plant cells,
   (c) screening progeny seeds from said fertile soybean plants for the desired levels of stearic acid, and
   (d) crushing said progeny seed to obtain said soybean oil containing higher-than-normal levels of stearic acid.

8. A method of producing oils from plant seed containing lower-than-normal levels of stearic acid comprising:
   (a) transforming a plant cell of an oil producing species with a chimeric gene of Claims 5 or 6,
   (b) growing sexually mature plants from said transformed plant cells of an oil producing species,
   (c) screening progeny seeds from said fertile plants for the desired levels of stearic acid, and
   (d) crushing said progeny seed to obtain said oil containing lower-than-normal levels of stearic acid.

9. A method of Claim 8 wherein said plant cell of an oil producing species is selected from the group consisting of soybean, rapeseed, sunflower, cotton, cocoa, peanut, safflower, and corn.

10. A method of Claim 7 wherein said step of transforming is accomplished by a process selected from the group consisting of Agrobacterium infection, electroporation, and high-velocity ballistic bombardment.

11. A method of Claim 8 wherein said step of transforming is accomplished by a process selected from the group consisting of Agrobacterium infection, electroporation, and high-velocity ballistic bombardment.

12. A method of producing mature soybean seed stearoyl-ACP desaturase enzyme in microorganisms comprising:
    (a) transforming a microorganism with a chimeric gene of Claim 6,
    (b) growing said transformed microorganism to produce quantities of said mature soybean seed stearoyl-ACP desaturase enzyme, and
    (c) isolating and purifying said mature soybean seed stearoyl-ACP desaturase enzyme.

13. A method of breeding soybean plants producing altered stearic acid levels in seed oil due to altered levels of stearoyl-ACP desaturase in said soybean plants by RFLP mapping comprising:
    (a) making a cross between two soybean varieties differing in stearic acid levels due to altered levels of stearoyl-ACP desaturase;
    (b) making a Southern blot of genomic DNA isolated from several progeny plants resulting from the cross following digestion with a suitable restriction enzyme that reveals polymorphism linked to the altered levels of stearic acid using a radiolabelled nucleic acid fragment of Claim 1 as a hybridization probe;
    (c) hybridizing the Southern blot with the radiolabelled nucleic acid fragment of Claim 1; and
    (d) selecting said soybean plants that inherit the RFLP linked to the desired level of stearic acid.

**Patentansprüche**

1. Isoliertes Nukleinsäurefragment, umfassend eine Nukleotidsequenz, die für Sojabohnensamen-Stearoyl-ACP-Desaturase kodiert, die den Nukleotiden 1 - 2243 in SEQ ID NO:1 entspricht, oder ein Sojabohnen-Nukleinsäurefragment, das im wesentlichen dazu homolog ist, das für eine funktionelle Stearoyl-ACP-Desaturase kodiert.

2. Isoliertes Nukleinsäurefragment nach Anspruch 1, worin die genannte Nukleotidsequenz für die Sojabohnensamen-Stearoyl-ACP-Desaturase-Vorstufe, entsprechend den Nukleotiden 70 - 1245 in SEQ ID NO:1, kodiert, oder ein Sojabohnen-Nukleinsäurefragment, das im wesentlichen dazu homolog ist und für eine funktionelle Stearoyl-ACP-Desaturase-Vorstufe kodiert.

3. Nukleinsäurefragment nach Anspruch 2, bei dem die genannte Nukleotidsequenz für das Stearoyl-ACP-Desaturase-Enzym von reifem Sojabohnensamen kodiert, das den Nukleotiden 166 - 1245 in SEQ ID NO:1 entspricht.

4. Chimäres Gen, das in der Lage ist, eine Sojabohnen-Pflanzenzelle zu transformieren, umfassend ein Nukleinsäurefragment nach Anspruch 1, das zweckorientiert mit geeigneten regulatorischen Sequenzen verknüpft ist, die eine Antisinn-Hemmung der Sojabohnensamen-Stearoyl-ACP-Desaturase in dem Samen erzeugen.

5. Chimäres Gen, das in der Lage ist, eine Pflanzenzelle einer ölproduzierenden Spezies zu transformieren, umfassend ein Nukleinsäurefragment nach Anspruch 2, das mit geeigneten regulatorischen Sequenzen zweckorientiert verknüpft ist, was zu einer Überexpression der genannten Sojabohnensamen-Stearoyl-ACP-Desaturase in dem Plastid der genannten Pflanzenzelle führt.

6. Chimäres Gen, das in der Lage ist, eine Pflanzenzelle einer ölproduzierenden Spezies zu transformieren, umfassend ein Nukleinsäurefragment nach Anspruch 3, das mit geeigneten regulatorischen Sequenzen zweckorientiert verknüpft ist, was zu der Expression des genannten Stearoyl-ACP-Desaturase-Enzyms von reifem Sojabohnensamen in dem Cytoplasma der genannten Pflanzenzelle führt.

7. Verfahren zur Herstellung von Sojabohnensamenöl, enthaltend höhere als normale Konzentrationen an Stearinsäure, umfassend:
    (a) Transformieren einer Sojabohnen-Pflanzenzelle mit einem chimären Gen nach Anspruch 4,
    (b) Züchten der fruchtbaren Sojabohnenpflanzen aus den genannten transformierten Sojabohnen-Pflanzenzellen,

(c) Überprüfung der zeugungsfähigen Samen aus den genannten fruchtbaren Sojabohnenpflanzen auf die gewünschten Stearinsäure-Konzentrationen und
(d) Zerquetschen des genannten zeugungsfähigen Samens, um das genannten Sojabohnenöl zu erhalten, das höhere als normale Stearinsäure-Konzentrationen enthält.

8. Verfahren zur Herstellung von Ölen aus Pflanzensamen, die niedrigere als normale Stearinsäure-Konzentrationen enthalten, umfassend:
(a) Transformieren einer Pflanzenzelle einer ölproduzierenden Spezies mit einem chimären Gen nach den Ansprüchen 5 oder 6,
(b) Züchten von sexuell reifen Pflanzen aus den genannten transformierten Pflanzenzellen einer ölproduzierenden Spezies,
(c) Überprüfung der zeugungsfähigen Samen aus den genannten fruchtbaren Pflanzen auf die gewünschten Stearinsäure-Konzentrationen, und
(d) Zerquetschen des genannten zeugungsfähigen Samens, um das genannte Öl zu erhalten, das niedrigere als normale Stearinsäure-Konzentrationen enthält.

9. Verfahren nach Anspruch 8, bei dem die genannte Pflanzenzelle einer ölproduzierenden Spezies aus der Gruppe ausgewählt wird, bestehend aus Sojabohne, Rapssamen, Sonnenblume, Baumwolle, Kakao, Erdnuß, Färberdistel und Mais.

10. Verfahren nach Anspruch 7, bei dem die genannte Stufe der Transformation durch ein Verfahren durchgeführt wird, ausgewählt aus der Gruppe, bestehend aus einer Agrobacterium-Infektion, einer Elektroporation und einer Hochgeschwindigkeitsstoßbombardierung.

11. Verfahren nach Anspruch 8, bei dem die genannte Transformationsstufe durch ein Verfahren durchgeführt wird, ausgewählt aus der Gruppe, bestehend aus einer Agrobacterium-Infektion, einer Elektroporation und einer Hochgeschwindigkeitsstoßbombardierung.

12. Verfahren zur Herstellung des Stearoyl-ACP-Desaturase-Enzyms von reifem Sojabohnensamen in Mikroorganismen, umfassend:
(a) Transformieren eines Mikroorganismus mit einem chimären Gen nach Anspruch 6,
(b) Züchten des genannten transformierten Mikroorganismus, um Mengen des genannten Stearoyl-ACP-Desaturase-Enzyms von reifem Sojabohnensamen zu produzieren und
(c) Isolieren und Reinigen des genannten Stearoyl-ACP-Desaturase-Enzyms von reifem Sojabohnensamen.

13. Verfahren zur Züchtung von Sojabohnenpflanzen, die aufgrund veränderter Konzentrationen der Stearoyl-ACP-Desaturase in den genannten Sojabohnenpflanzen durch RFLP-Kartierung veränderte Stearinsäure-Konzentrationen in dem Samenöl produzieren, umfassend:
(a) Kreuzen zweier Sojabohnen-Varietäten, die sich aufgrund der veränderten Konzentrationen der Stearoyl-ACP-Desaturase in den Stearinsäure-Konzentrationen unterscheiden,
(b) Anfertigen eines Southern Blots der genomischen DNA, die aus mehreren, aus der Kreuzung hervorgegangenen zeugungsfähigen Pflanzen isoliert worden ist, und anschließender Verdau mit einem geeigneten Restriktionsenzym, das den Polymorphismus, der mit den veränderten Stearinsäure-Konzentrationen verknüpft ist, aufdeckt, wobei ein radioaktiv markiertes Nukleinsäurefragment nach Anspruch 1 als Hybridisierungssonde verwendet wird,
(c) Hybridisierung des Southern Blots mit dem radioaktiv markierten Nukleinsäurefragment nach Anspruch 1, und
(d) Auswählen der genannten Sojabohnenpflanzen, die das RFLP, das mit der gewünschten Stearinsäure-Konzentration verknüpft ist, vererben.

**Revendications**

1. Un fragment d'acide nucléique isolé comprenant une séquence nucléotidique codant pour la stéaroyl-ACP-désaturase de graine de soja correspondant aux nucléotides 1 à 2243 de SEQ ID N° 1, ou tout fragment d'acide nucléique de soja sensiblement homologue à celui-ci codant pour une stéaroyl-ACP-désaturase fonctionnelle.

2. Un fragment d'acide nucléique isolé de la revendication 1, dans lequel ladite séquence nucléotidique code pour le précurseur de stéaroyl-ACP-désaturase de graine de soja correspondant aux nucléotides 70 à 1245 de SEQ ID N° 1, ou tout fragment d'acide nucléique de soja sensiblement homologue à celui-ci codant pour un précurseur de stéaroyl-ACP-désaturase fonctionnel.

3. Un fragment d'acide nucléique de la revendication 2, dans lequel ladite séquence nucléotidique code pour la stéaroyl-ACP-désaturase de graine de soja mûre, correspondant aux nucléotides 166 à 1245 de SEQ ID N° 1.

4. Un gène chimérique capable de transformer une cellule de soja, comprenant un fragment d'acide nucléique de la revendication 1, lié fonctionnellement à des séquences régulatrices appropriées produisant une inhibition anti-sens de la stéaroyl-ACP-désaturase dans la graine.

5. Un gène chimérique capable de transformer une cellule végétale d'une espèce productrice d'huile, comprenant un fragment d'acide nucléique de la revendication 2 lié fonctionnellement à des séquences régulatrices appropriées donnant lieu à une surexpression de ladite stéaroyl-ACP-désaturase de graine de soja dans le plastide de ladite cellule végétale.

6. Un gène chimérique capable de transformer une cellule végétale d'une espèce productrice d'huile, comprenant un fragment d'acide nucléique de la revendication 3 lié fonctionnellement à des séquences régulatrices appropriées donnant lieu à l'expression de ladite stéaroyl-ACP-désaturase de graine de soja mûre dans le cytoplasme de ladite cellule végétale.

7. Un procédé de production d'huile de graine de soja contenant des taux supérieurs à la normale d'acide stéarique, consistant à :
   (a) transformer une cellule de soja avec un gène chimérique de la revendication 4,
   (b) faire croître des plants de soja fertiles à partir de cellules de soja transformées,
   (c) sélectionner des graines de descendance provenant desdits plants de soja fertiles pour les taux souhaités d'acide stéarique, et
   (d) broyer lesdites graines de descendance pour obtenir ladite huile de soja contenant des taux d'acide stéarique supérieurs à la normale.

8. Un procédé de production d'huiles à partir de graines végétales contenant des taux d'acide stéarique inférieurs à la normale, consistant à :
   (a) transformer une cellule végétale d'une espèce productrice d'huile avec un gène chimérique de la revendication 5 ou 6,
   (b) faire croître des plants sexuellement matures à partir desdites cellules végétales transformées d'une espèce productrice d'huile,
   (c) sélectionner des graines de descendance provenant desdits plants fertiles pour les taux désirés d'acide stéarique, et
   (d) broyer lesdites graines de descendance pour obtenir l'huile contenant des taux d'acide stéarique inférieurs à la normale.

9. Un procédé de la revendication 8, dans lequel ladite cellule végétale d'une espèce productrice d'huile est choisie dans le groupe formé par le soja, le colza, le tournesol, le cotonnier, le cacaoyer, l'arachide, le carthame et le maïs.

10. Un procédé de la revendication 7, dans lequel ladite étape de transformation est exécutée par un procédé choisi dans le groupe formé par une infection par *Agrobacterium*, une électroporation et un bombardement balistique à grande vitesse.

11. Un procédé de la revendication 8, dans lequel ladite étape de transformation est exécutée par un procédé choisi dans le groupe formé par une infection par *Agrobacterium*, une électroporation et un bombardement balistique à grande vitesse.

12. Un procédé de production de stéaroyl-ACP-désaturase de graine de soja mûre dans des microorganismes, consistant à :
   (a) transformer un microorganisme avec un gène chimérique de la revendication 6,

(b) faire croître ledit microorganisme transformé pour produire des quantités de stéaroyl-ACP-désaturase de graine de soja mûre, et
(c) isoler et purifier ladite stéaroyl-ACP-désaturase de graine de soja mûre.

13. Un procédé de culture de plants de soja produisant des taux altérés d'acide stéarique dans l'huile de graines à cause de taux altérés de stéaroyl-ACP-désaturase dans lesdits plants de soja par cartographie de polymorphisme en longueur des fragments de restriction (RFLP), consistant à :
(a) effectuer un croisement entre deux variétés de soja différant par les taux d'acide stéarique à cause de taux altérés de stéaroyl-ACP-désaturase ;
(b) effectuer une analyse Southern blot d'ADN génomique isolé de plusieurs plants de descendance résultant du croisement, après digestion avec une enzyme de restriction appropriée qui révèle un polymorphisme lié aux taux altérés d'acide stéarique en utilisant un fragment d'acide nucléique de la revendication 1 radiomarqué comme sonde d'hybridation ;
(c) hybrider le Southern blot avec le fragment d'acide nucléique de la revendication 1 radiomarqué; et
(d) sélectionner lesdits plants de soja qui héritent du RFLP lié au taux désiré d'acide stéarique.

# EXHIBIT H

**PATENT**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:
Thomas R. ADAMS et al.

Serial No.: 08/113,561

Filed: August 25, 1995

For: METHOD AND COMPOSITIONS FOR THE PRODUCTION OF STABLY TRANSFORMED, FERTILE MONOCOT PLANTS AND CELLS THEREOF

Group Art Unit: 1638

Examiner: Fox, David T.

Atty. Dkt. No.: DEKM:055US

**DECLARATION OF VIRGINIA URSIN UNDER 37 C.F.R. § 1.132**

Assistant Commissioner for Patents
Washington, D.C. 20231

I, VIRGINIA URSIN HEREBY DECLARE AS FOLLOWS:

1. I have been employed by Calgene Inc. and Monsanto Company since 1989, currently with the position of Project Lead Lipid Technologies. Monsanto Company is the parent company of wholly owned subsidiaries Calgene Inc. and Monsanto Company.

2. I hold a Ph.D. in Genetics from University of California. I have been conducting research in the area of agricultural biotechnology since 1987.

3. I understand that the Patent and Trademark Office Examiner in charge of assessing the patentability of the referenced patent application has rejected the claims as not being supported by adequate information in the specification to show that introduction of a heterologous fatty acid desaturase in maize would result in an altered grain composition trait that would render such a maize plant identifiable over the corresponding untransformed maize plants which do not comprise the heterologous gene.

4. Therefore, I am providing the present Declaration to submit further data that demonstrates that heterologous expression of fatty acid desaturase genes in maize alters the fatty acid profile of transgenic plants in a predictable and consistent manner that renders them identifiable over corresponding non-transgenic plants.

5. *Vector Construction and Transformation of Maize*

A binary vector was constructed to express a Δ15-desaturase and a Δ6-desaturase in maize embryo and aleurone tissue. This construct was prepared with the globulin promoter (see, e.g., Table 3, Regulatory Sequence 123 of Pat. Appl. Serial no. 08/113,561) driving expression of a mutagenized *Neurospora crassa* Δ15 desaturase and a *Mortierella alpina* Δ6 desaturase (SEQ ID NO:21, bp 71-1444) (U.S. Pat. No. 6,075,183). The *M. alpina* Δ6 desaturase was cloned into a globulin expression cassette shuttle vector, pMON67624, resulting in pMON82809. The mutagenized *N. crassa* Δ15 desaturase was cloned into a globulin expression cassette vector, pMON67624, resulting in pMON82810.

The two globulin desaturase expression cassettes were then cloned into the pMON30167 1T maize binary vector containing the CP4 marker gene for glyphosate resistance. The first expression cassette containing the *M. alpina* Δ6 desaturase was cloned into pMON30167, resulting in pMON82811. The second expression cassette containing the mutagenized *N. crassa* Δ15 desaturase was then cloned into pMON82811, resulting in a maize transformation construct designated pMON82812. The resulting vector was introduced into maize via *Agrobacterium tumefaciens*-mediated transformation as known to one of skill in the art, *e.g.*, U.S. Patent Nos. 5,591,616 and 6,603,061.

6. *Fatty Acid Analysis*

The fatty acid composition of single immature kernels of plants transformed with vector pMON82812 was determined by lyophilizing maize kernels and extracting the kernels with toluene and 5.0 % (wt/vol) sulfuric acid in methanol, followed by heat treatment. Following the heat treatment, the reaction mixture was extracted with heptane followed by aqueous sodium chloride (10% wt/vol). After partitioning at room temperature, the organic phase was analyzed by GLC (Hewlett Packard model 6890 (120volt) equipped with a split/splitless capillary inlet (250°C) and a flame ionization detector (270°C). The column was a Supelco 24077 (0.25 mm

od. x 15 m length) with a 0.25 µm bonded polyethylene glycol stationary phase. The fatty acid methyl esters are identified by retention time comparison to commercial standards. Qualitative weight percent compositions are calculated as area percents of identified peaks.

The data in Table 1 below demonstrate fatty acid profiles for kernels of transgenic maize expressing the mutagenized *Neurospora crassa* Δ15 desaturase and *Mortierella alpina* Δ6 desaturase of pMON82812 and accumulating SDA (18:4) and GLA (18:3), which are not seen in untransformed lines. These events also demonstrate increased accumulation of ALA (18:3) and decreased accumulation of LA (18:2). Of 180 seeds tested, 67 contained SDA and GLA, 25 contained GLA but not SDA, and 88 were wild type with respect to GLA and SDA.

**TABLE 1: Fatty Acid Analysis of Single Immature Maize Kernels Expressing SDA and/or GLA**

| Pedigree | Event | Gen | Oleic (18:1) | LA (18:2) | GLA (18:3) | ALA (18:3) | SDA (18:4) |
|---|---|---|---|---|---|---|---|
| ZM_S103121:@. | ZM_S103121 | R1 | 21.5 | 25.54 | 1.22 | 28.45 | 2.08 |
| ZM_S103121:@. | ZM_S103121 | R1 | 21.21 | 29.49 | 1.32 | 24.92 | 1.83 |
| ZM_S103121:@. | ZM_S103121 | R1 | 18.93 | 32.23 | 1.85 | 24.58 | 1.62 |
| ZM_S103121:@. | ZM_S103121 | R1 | 19.99 | 29.81 | 1.46 | 26.36 | 1.36 |
| ZM_S103435/LH244 | ZM_S103435 | F1 | 19.56 | 34.49 | 0.7 | 23.78 | 0.88 |
| ZM_S103121:@. | ZM_S103121 | R1 | 17.24 | 35.26 | 1.3 | 23.99 | 0.8 |
| ZM_S103435/LH244 | ZM_S103435 | F1 | 19.78 | 36.27 | 0.61 | 22.2 | 0.6 |
| ZM_S103435/LH244 | ZM_S103435 | F1 | 19.61 | 34.73 | 0.56 | 23.59 | 0.58 |
| ZM_S103432:@. | ZM_S103432 | R1 | 19.44 | 33.05 | 0.83 | 25.89 | 0.57 |
| ZM_S103121:@. | ZM_S103121 | R1 | 18.99 | 33.41 | 0.64 | 25.39 | 0.55 |
| ZM_S103432:@. | ZM_S103432 | R1 | 19.72 | 33.14 | 0.88 | 24.82 | 0.52 |
| ZM_S103121:@. | ZM_S103121 | R1 | 18.08 | 35.7 | 0.97 | 22.92 | 0.49 |
| ZM_S103435/LH244 | ZM_S103435 | F1 | 19.34 | 35.22 | 0.54 | 23.56 | 0.47 |
| ZM_S103435/LH244 | ZM_S103435 | F1 | 17.95 | 36.96 | 0.54 | 22.88 | 0.43 |
| ZM_S103433/LH244 | ZM_S103433 | F1 | 19.18 | 43.84 | 0.22 | 14.98 | 0.36 |
| ZM_S103437/LH244 | ZM_S103437 | F1 | 20.26 | 42.75 | 0.69 | 14.73 | 0.36 |
| ZM_S103435/LH244 | ZM_S103435 | F1 | 19.75 | 41.59 | 0.58 | 17.1 | 0.35 |
| ZM_S103437/LH244 | ZM_S103437 | F1 | 21.04 | 42.89 | 0.67 | 14.03 | 0.35 |
| ZM_S103110:@. | ZM_S103110 | R1 | 18.62 | 37.51 | 0.57 | 22.51 | 0.34 |
| ZM_S103432:@. | ZM_S103432 | R1 | 19.5 | 38.38 | 0.77 | 20.26 | 0.34 |

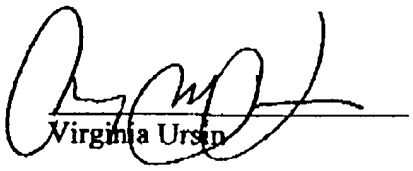
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ZM_S103437/LH244 | ZM_S103437 | F1 | 19.18 | 44.68 | 0.64 | 14.1 | 0.32 |
| ZM_S103435/LH244 | ZM_S103435 | F1 | 19.11 | 38.92 | 0.38 | 20.6 | 0.3 |
| ZM_S103432:@. | ZM_S103432 | R1 | 17.99 | 40.62 | 0.74 | 19.34 | 0.3 |
| ZM_S103432:@. | ZM_S103432 | R1 | 17.93 | 40.5 | 0.78 | 19.26 | 0.3 |
| ZM_S103432:@. | ZM_S103432 | R1 | 19.55 | 38.86 | 0.75 | 19.66 | 0.29 |
| ZM_S103110:@. | ZM_S103110 | R1 | 19.84 | 48.95 | 0.52 | 9.9 | 0.28 |
| ZM_S103110:@. | ZM_S103110 | R1 | 19.21 | 37.14 | 0.41 | 22.57 | 0.28 |
| ZM_S103432:@. | ZM_S103432 | R1 | 19.33 | 38.28 | 0.72 | 20.5 | 0.28 |
| ZM_S103435/LH244 | ZM_S103435 | F1 | 19.68 | 42.85 | 0.59 | 15.62 | 0.28 |
| LH244/ZM_S103123 | ZM_S103123 | F1 | 18.01 | 40.3 | 0.36 | 20.22 | 0.27 |
| ZM_S103110:@. | ZM_S103110 | R1 | 19.51 | 42.17 | 0.54 | 17.37 | 0.27 |
| ZM_S103437/LH244 | ZM_S103437 | F1 | 20.08 | 45.06 | 0.58 | 12.99 | 0.26 |
| ZM_S103435/LH244 | ZM_S103435 | F1 | 19.5 | 42.99 | 0.45 | 15.9 | 0.25 |
| LH244/ZM_S103123 | ZM_S103123 | F1 | 19.41 | 43.27 | 0.41 | 15.9 | 0.25 |
| ZM_S103168/LH244 | ZM_S103168 | F1 | 18.78 | 44.78 | 1.32 | 14.48 | 0.24 |
| ZM_S103110:@. | ZM_S103110 | R1 | 17.95 | 41.13 | 0.58 | 19.46 | 0.23 |
| ZM_S103432:@. | ZM_S103432 | R1 | 17.81 | 40.06 | 0.67 | 20.1 | 0.22 |
| ZM_S103435/LH244 | ZM_S103435 | F1 | 19.26 | 44.01 | 0.51 | 14.7 | 0.22 |
| ZM_S103436/LH244 | ZM_S103436 | F1 | 19.88 | 45.18 | 0.49 | 13.4 | 0.22 |
| ZM_S103168/LH244 | ZM_S103168 | F1 | 20.19 | 43.05 | 1.06 | 15.11 | 0.21 |
| ZM_S103110:@. | ZM_S103110 | R1 | 19 | 46.61 | 0.46 | 13.13 | 0.21 |
| LH244/ZM_S103431 | ZM_S103431 | F1 | 18.73 | 37.53 | 0.27 | 22.93 | 0.2 |
| ZM_S103099/LH244 | ZM_S103099 | F1 | 19.32 | 38.76 | 0.54 | 20.49 | 0.2 |
| LH244/ZM_S103123 | ZM_S103123 | F1 | 20.26 | 36.51 | 0.28 | 22.25 | 0.19 |
| ZM_S103433/LH244 | ZM_S103433 | F1 | 19.4 | 43.5 | 0.21 | 15.38 | 0.19 |
| ZM_S103168/LH244 | ZM_S103168 | F1 | 19.44 | 44.9 | 0.99 | 13.94 | 0.19 |
| ZM_S103168/LH244 | ZM_S103168 | F1 | 20.59 | 44.07 | 0.79 | 13.9 | 0.19 |
| ZM_S103110:@. | ZM_S103110 | R1 | 19 | 46.29 | 0.51 | 13.61 | 0.19 |
| ZM_S103437/LH244 | ZM_S103437 | F1 | 19.9 | 45.07 | 0.51 | 13.25 | 0.19 |
| ZM_S103436/LH244 | ZM_S103436 | F1 | 19.97 | 45.59 | 0.41 | 13.06 | 0.19 |
| ZM_S103168/LH244 | ZM_S103168 | F1 | 20.92 | 49.2 | 0.58 | 9.12 | 0.18 |
| LH244/ZM_S103431 | ZM_S103431 | F1 | 18.27 | 37.66 | 0.32 | 22.96 | 0.18 |
| ZM_S103103/LH244 | ZM_S103103 | F1 | 19.19 | 46.83 | 0.7 | 12.21 | 0.18 |
| ZM_S103436/LH244 | ZM_S103436 | F1 | 18.34 | 48.08 | 0.37 | 11.98 | 0.18 |
| ZM_S103433/LH244 | ZM_S103433 | F1 | 19.71 | 43.32 | 0.22 | 15.12 | 0.17 |
| ZM_S103436/LH244 | ZM_S103436 | F1 | 19.16 | 47.28 | 0.37 | 11.95 | 0.17 |
| LH244/ZM_S103431 | ZM_S103431 | F1 | 18.78 | 37.05 | 0.27 | 23.34 | 0.16 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ZM_S103099/LH244 | ZM_S103099 | F1 | 18.66 | 40.7 | 0.45 | 19.85 | 0.16 |
| ZM_S103433/LH244 | ZM_S103433 | F1 | 19.78 | 43.09 | 0.22 | 15.39 | 0.16 |
| ZM_S103430/LH244 | ZM_S103430 | F1 | 19.92 | 42.81 | 0.69 | 16.07 | 0.15 |
| ZM_S103437/LH244 | ZM_S103437 | F1 | 20.17 | 46.64 | 0.48 | 10.88 | 0.15 |
| LH244/ZM_S103123 | ZM_S103123 | F1 | 19.87 | 40.51 | 0.23 | 18.3 | 0.14 |
| ZM_S103430/LH244 | ZM_S103430 | F1 | 19.25 | 43.68 | 0.52 | 15.4 | 0.13 |
| ZM_S103436/LH244 | ZM_S103436 | F1 | 19.49 | 47.11 | 0.31 | 11.99 | 0.12 |
| ZM_S103103/LH244 | ZM_S103103 | F1 | 19.77 | 47.5 | 0.37 | 11.33 | 0.12 |
| ZM_S103432:@. | ZM_S103432 | R1 | 18.68 | 42.32 | 0.42 | 17.36 | 0.11 |
| ZM_S103103/LH244 | ZM_S103103 | F1 | 19.85 | 47.51 | 0.35 | 11.33 | 0.1 |
| ZM_S103168/LH244 | ZM_S103168 | F1 | 17.33 | 51.87 | 1.13 | 7.55 | 0 |
| ZM_S103099/LH244 | ZM_S103099 | F1 | 18.41 | 40.56 | 0.67 | 19.94 | 0 |
| ZM_S103433/LH244 | ZM_S103433 | F1 | 17.88 | 52.22 | 0.62 | 7.58 | 0 |
| ZM_S103097/LH244 | ZM_S103097 | F1 | 18.77 | 47.41 | 0.61 | 12.28 | 0 |
| ZM_S103430/LH244 | ZM_S103430 | F1 | 18.36 | 46.05 | 0.57 | 13.91 | 0 |
| ZM_S103110:@. | ZM_S103110 | R1 | 18.48 | 45.05 | 0.48 | 15.43 | 0 |
| ZM_S103436/LH244 | ZM_S103436 | F1 | 19.8 | 46.28 | 0.42 | 13.34 | 0 |
| ZM_S103103/LH244 | ZM_S103103 | F1 | 20.11 | 47.5 | 0.42 | 11 | 0 |
| ZM_S103099/LH244 | ZM_S103099 | F1 | 18.38 | 41.21 | 0.42 | 19.45 | 0 |
| ZM_S103099/LH244 | ZM_S103099 | F1 | 18.62 | 41.08 | 0.4 | 19.14 | 0 |
| ZM_S103103/LH244 | ZM_S103103 | F1 | 19.9 | 48.09 | 0.38 | 10.89 | 0 |
| LH244/ZM_S103123 | ZM_S103123 | F1 | 18.47 | 45.81 | 0.37 | 14.52 | 0 |
| ZM_S103103/LH244 | ZM_S103103 | F1 | 19.8 | 48.52 | 0.32 | 10.36 | 0 |
| LH244/ZM_S103123 | ZM_S103123 | F1 | 18.45 | 46.48 | 0.32 | 13.79 | 0 |
| LH244/ZM_S103123 | ZM_S103123 | F1 | 19.27 | 41.97 | 0.24 | 17.97 | 0 |
| ZM_S103433/LH244 | ZM_S103433 | F1 | 19.84 | 43.54 | 0.23 | 14.61 | 0 |
| ZM_S103433/LH244 | ZM_S103433 | F1 | 19.68 | 43.73 | 0.22 | 15.15 | 0 |
| ZM_S103433/LH244 | ZM_S103433 | F1 | 18.85 | 43.9 | 0.22 | 15.03 | 0 |
| ZM_S103433/LH244 | ZM_S103433 | F1 | 19.72 | 44.98 | 0.19 | 13.42 | 0 |
| ZM_S103434/LH244 | ZM_S103434 | F1 | 20.56 | 43.62 | 0.14 | 15.03 | 0 |
| ZM_S103434/LH244 | ZM_S103434 | F1 | 19.8 | 44.48 | 0.14 | 14.84 | 0 |
| ZM_S103434/LH244 | ZM_S103434 | F1 | 19.85 | 44.8 | 0.13 | 14.68 | 0 |
| ZM_S103434/LH244 | ZM_S103434 | F1 | 18.53 | 45.5 | 0.12 | 14.86 | 0 |
| ZM_S103434/LH244 | ZM_S103434 | F1 | 20.17 | 44.8 | 0.12 | 14.43 | 0 |
| ZM_S103434/LH244 | ZM_S103434 | F1 | 19.38 | 45.86 | 0.12 | 13.73 | 0 |

7. The results of the above studies demonstrated that expression of a fatty acid desaturase gene in maize alters the fatty acid profile in a manner that renders the transgenic plants identifiable over corresponding non-transgenic plants. The results further confirm that the alteration of fatty acid profiles in maize occurs in a predictable manner that is consistent with the enzymatic activity of the fatty acid desaturase that is introduced into a given maize plant.

8. I hereby declare that all statements made herein of my knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

13 October, 2004
Date


Virginia Ursin

# This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☑ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

**IMAGES ARE BEST AVAILABLE COPY.**
**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**